NONRELATIVISTIC QUANTUM MECHANICS

Third Edition

Anton Z. Capri

World Scientific

NONRELATIVISTIC QUANTUM MECHANICS

Third Edition

Anton Z. Capri
Department of Physics
University of Alberta, Canada

World Scientific
Singapore • New Jersey • London • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.
P O Box 128, Farrer Road, Singapore 912805
USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data
Capri, Anton Z.
Nonrelativistic quantum mechanics / Anton Z. Capri. -- 3rd ed.
p. cm.
Includes bibliographical references and index.
ISBN 981024634X (alk. paper) -- ISBN 981024651X (pbk. : alk. paper)
1. Nonrelativistic quantum mechanics. 2. Perturbation (Quantum dynamics) I. Title.

QC174.24.N64 C37 2002
530.12--dc21 2002028870

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

This book is printed on acid-free paper.

Printed in Singapore by Uto-Print

To Kim, Karin, Irene and Skaidrite
for giving me time and love.

Preface to the First and Second Editions

Most textbooks start as a set of lecture notes which expand and undergo numerous revisions over the years. The present version is no exception and is the culmination of many revisions of such lecture notes. I have taught quantum mechanics at several different levels and thus attempted to make this book of sufficiently broad scope for different courses. The book includes material for undergraduate courses as well as a one semester graduate course. As a guide for possible division of the material in this book I have included a few pages in the following section entitled "How to use this Book".

It is a long way from lecture notes to a text-book and on this journey I have been greatly helped by many people. My first debt is to my students who forced me to clarify and expand my lecture notes. Without their prodding this book would never have been started.

Dr. M. Razavy has also been most generous with his time and ideas. Not only did he test this book in his quantum mechanics classes, but he willingly proof-read all of the present version. He also helped with many pertinent references and comments; his criticism of chapter 21 was particularly useful. Without his constant encouragement and support my enthusiasm would have flagged many times and this book might never have been written.

During the early stages, Dr. W. Brouwer also helped considerably with his detailed criticisms of the first fourteen chapters. I thank him for his generous contribution of time and ideas.

The first version of this book was expertly typed by Mrs. M. Yiu. Had she not been here to read my scrawl, I am sure this book would never been written. I thank her for her patience and tolerance.

The diagrams are due to the heroic efforts of Mrs. J. Hube. Her attention to detail turned some rather skimpy sketches into actual diagrams. I am most grateful for her skills.

Finally, I acknowledge with many thanks the expertly accomplished job of preparing the final manuscript. This was ably handled by Christine Fischer and Laura Heiland. Their friendly attitude did much to make a painful job easier.

I would also like to thank all of my colleagues who, in one way or another, influenced me during our coffee-room discussions. Their contributions, although less tangible, are nevertheless very real. It is my hope that in writing this book I may have helped a few students to discover the beauty of quantum mechanics.

Anton Z. Capri
Edmonton, Alberta
April, 1985

Preface to the Third Edition

In this new edition I have added some topics that have aroused interest in the last two decades. I have also omitted or shortened the discussion of certain topics that are no longer of major interest. The presentation of the ideas of distribution theory and rigged Hilbert spaces has been well received by my students and remains virtually unchanged. Even though these concepts do not immediately empower students to better compute physical effects, it seems to make them feel more comfortable with the whole machinery of quantum mechanics. As I like to tell my students, "A physicist has to learn rigorous mathematics so that he or she knows when it is safe to be sloppy".

A topic like quantum mechanics is always listed as part of "modern" physics. This, in spite of the fact that the subject is now more than three quarters of a century old. Nevertheless, quantum mechanics continues to produce fresh problems that keep the subject young and vibrant. In this category one has to view the Aharonov-Bohm effect, as well as Berry's phase. In the present edition I have included a discussion of both of these.

Quantum mechanics is a most robust theory and even after more than three quarters of a century there is not a single experiment in conflict with it. Yet, epistemological problems remain. To give some flavour of these problems I have included a short discussion of the recent models on decoherence since this promises to be an area of ongoing research, yet is accessible to the earnest student.

The number of problems at the end of each chapter has been expanded and revised in an attempt to make them more useful to the students. In certain cases the subject matter of the main text is amplified in the problem section.

Finally I would like to thank Professor Ruggero Ferrari who some years ago first introduced me to the questionable pleasures of using Teχ with all the backslashes and braces. Without his goading I might never have attempted to master that art and a revision of this book.

As in the first edition, it is still my hope that in some way I have contributed to helping students understand and enjoy this most wonderful subject.

Anton Z. Capri
Edmonton, Alberta
June, 2002.

Contents

Chapter 1

The Breakdown of Classical Mechanics

1.1 Introduction

During the nineteenth century many of the great advances in physics of the eighteenth century were consolidated and extended. In addition, the theory of electromagnetism was completed by J.C. Maxwell. Except for a few unexplained effects or anomalies there seemed little more in terms of fundamental physics to be done by the beginning of the twentieth century. Yet it is precisely in the year 1900 that quantum theory starts with Planck's formula for blackbody radiation. Soon there were a host of experimental results, both new ones and earlier ones that again attracted attention. All of these pointed to flaws in the physics of the nineteenth century. In almost all cases these anomalies resulted when Newtonian mechanics and electromagnetism were simultaneously involved. In trying to elucidate these various experimental facts a new theory of physics, quantum theory, was born.

In the next few sections we briefly examine several of these experiments and discuss them with some modern hindsight. First we consider blackbody radiation from the pre-quantum or classical point of view. We then turn to a consideration of the stability of the classical Rutherford atom. Although Rutherford had experimentally demonstrated the planet-like structure of the electrons in atoms, his model caused a lot of theoretical problems.

An even older effect, dating back to Hertz in 1887, the photoelectric effect, provides another example of the complete breakdown of classical physics. After considering this effect we discuss the elastic scattering of light off electrons, the Compton effect, which also clearly demonstrates the corpuscular nature of light. On the other hand, the Davisson-Germer experiments demonstrate very clearly the wavelike nature of electrons and are discussed in the following section. As a final example we consider the Franck-Hertz effect. This is a beautiful experiment demonstrating that only certain definite quanta of energy can be absorbed by atoms in inelastic collisions with electrons.

1.2 Blackbody Radiation

Blackbody radiation refers to the equilibrium radiant energy to be found inside a cavity whose walls are completely opaque and held at a fixed temperature T. Such a cavity is called a blackbody cavity. The interest in such radiation derives from the fact that this radiation is independent of the nature of the walls of the cavity (their material properties, or their geometry); the spectral properties of the radiation depend only on the temperature of the walls. A very simple proof of this fact, utilizing only the second law of thermodynamics, was given by Kirchhoff [1.1]. The fact that the radiation depends only on the temperature of the walls means that it is somehow universal. Most glowing bodies such as a hot piece of iron or our sun are good approximations of a blackbody. By measuring the spectrum of their radiant energy we can determine their temperature. Since blackbody radiation was so simple, a theory for its spectrum was soon derived from classical mechanics and electromagnetism. The resultant spectral formula, called the *Rayleigh-Jeans Law* proved to fail completely at high frequencies. We now derive this classical formula.

Completely opaque walls for a cavity can be described mathematically by assuming that all the radiation is reflected at the walls. Thus, we want standing waves inside the cavity. If the cavity is a cube with sides of length L, then the components of the wavelength of the radiation in each direction must exactly fit into L. This means that there are nodes at the walls. Thus, we have for the component of the wavelength of the lowest frequency mode in the x-direction.

$$\lambda_{x,1} = \frac{2L}{1} \; . \tag{1.2.1}$$

The next mode has:

$$\lambda_{x,2} = \frac{2L}{2} \; . \tag{1.2.2}$$

The 3rd mode has:

$$\lambda_{x,3} = \frac{2L}{3} \tag{1.2.3}$$

and so on. The nth mode has

$$\lambda_{x,n} = \frac{2L}{n_x} \; . \tag{1.2.4}$$

$$n_x = 1, 2, 3, \ldots \; . \tag{1.2.5}$$

A similar discussion holds for the y and z directions. It is useful to convert these formulas into formulas involving wave numbers because then the integers

n_x, n_y, n_z appear in the numerator. The wave numbers k_x, k_y, k_z are related to the corresponding wavelengths by

$$k_x = \frac{2\pi}{\lambda_x}\ ,\ k_y = \frac{2\pi}{\lambda_y}\ ,\ k_z = \frac{2\pi}{\lambda_z}\,. \tag{1.2.6}$$

The different vibration modes are therefore characterized by three integers (n_x, n_y, n_z) giving the total wave number

$$k = \sqrt{{k_x}^2 + {k_y}^2 + {k_z}^2}. \tag{1.2.7}$$

Using (1.2.4) and (1.2.6) we get

$$k^2 = \frac{(2\pi)^2}{4L^2}({n_x}^2 + {n_y}^2 + {n_z}^2). \tag{1.2.8}$$

The wave number k is also related to the frequency ν by

$$\nu = \frac{ck}{2\pi}. \tag{1.2.9}$$

where c is the speed of light.

Now each mode of vibration of the electromagnetic field can be considered as two degrees of freedom of the field; the different vibration modes are independent of each other. One degree of freedom is associated with the kinetic energy and one with the potential energy since the modes of the field are treated as harmonic oscillators. However, according to the *equipartition principle* of statistical mechanics we have, for a temperature T, an amount of energy $1/2\ k_BT$ for each degree of freedom of the field. Here T is the temperature of the cavity wall and k_B is Boltzmann's constant. From the equipartition principle we can therefore write the formula:

> The energy dU in a frequency interval between ν and $\nu + d\nu$
> $= k_BT \times$ (number of modes of oscillation in this interval).

So to obtain the blackbody spectrum requires that we count the number of modes of oscillation corresponding to a frequency interval between ν and $\nu + d\nu$. It is easier to first obtain all modes up to a given frequency ν. This is simply the number of points (the volume) inside one quadrant (i.e. 1/8) of a sphere (since n_x, n_y, n_z are all positive) whose equation according to equation (1.2.8) is:

$$ {n_x}^2 + {n_y}^2 + {n_z}^2 = \frac{4L^2}{c^2}\nu^2 \tag{1.2.10}$$

The result is

$$N = \frac{1}{8} \times \frac{4}{3}\pi\left(\frac{2L\nu}{c}\right)^3 \times 2\ . \tag{1.2.11}$$

The last factor of 2 is due to the fact that for light two independent polarizations for each vibration mode are possible. Thus, the number of degrees of freedom is increased by this factor of 2.

The number of modes dN in the frequency interval between ν and $\nu + d\nu$ is now given by:

$$dN = \pi \left(\frac{2L}{c} \right)^3 \nu^2 d\nu \tag{1.2.12}$$

Hence, using our formula for the energy per unit volume in this frequency interval we get:

$$du = \frac{1}{V} k_B T dN = \frac{8\pi k_B T}{c^3} \nu^2 d\nu \tag{1.2.13}$$

where we have used the fact that the volume $V = L^3$. So by applying the classical equipartition theorem and classical electromagnetic theory we obtain for the energy density the unambiguous result

$$\rho(\nu) = \frac{du}{d\nu} = \frac{8\pi k_B T}{c^3} \nu^2 \, . \tag{1.2.14}$$

For low frequencies this result agrees splendidly with the experimental spectrum, but for high frequencies, as shown in figure 1.1, this result fails miserably. Furthermore, on purely theoretical grounds alone, equation (1.2.14) must be wrong since it predicts that for a cavity of volume V at a temperature T the total radiant energy in this volume is given by

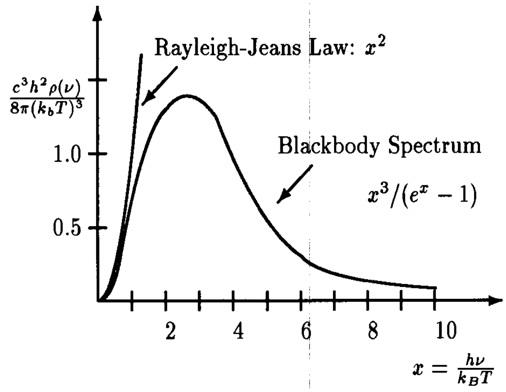

Figure 1.1: Blackbody radiation.

$$E = V \int_0^\infty \rho(\nu) d\nu = \frac{8\pi k_B T V}{c^3} \int_0^\infty \nu^2 d\nu \, . \tag{1.2.15}$$

Clearly this is an utterly nonsensical result. Our eyes are not seared by X-rays and gamma rays when we look at a cold stove or even a hot stove.

Exactly in the year 1900, Planck, who had previously found an analytic formula that fit the experimental results, gave a theoretical derivation of his formula. To "derive" his formula Planck applied the same tools as we did above, but in addition he made the very radical assumption that radiation of frequency ν carries energy

$$E = h\nu \tag{1.2.16}$$

where, as determined by experiment, it was found that $h = 6.63 \times 10^{-34}$ Joule-seconds and is now known as Planck's constant and plays a fundamental role in all of modern physics. Planck's energy-frequency relation equation (1.2.16) was a completely new and revolutionary equation in physics. Until Planck, the energy of a wave could be any number and was proportional to the square of the amplitude of the wave. Suddenly the energy was quantized in lumps proportional to the frequency. This was a very exciting new development, but it took a while to make an impact on the physics community. In section 1.10 we derive Planck's radiation formula by using his energy-frequency relation.

1.3 Stability of Atoms: Discrete Spectral Lines

In the early part of this century, atoms were beginning to be taken seriously by physicists as actual physical objects and not just models. As a consequence experiments to explore the structure of atoms were begun. The very reasonable model, due to J.J. Thomson, of an atom as a cloud of positive charge with bits of negative charge (the electrons) interspersed was proven wrong by experiment.

In a long series of brilliant scattering experiments Rutherford showed conclusively in 1911 that the atom consisted of a tiny positive core of magnitude about 10^{-12} cm, called the nucleus, with electrons whirling about this nucleus at a much greater distance of about 10^{-8} cm. This planetary model caused a crisis because atoms were known to be stable lasting for millions of years. But, as we now show, such planetary atoms are intrinsically unstable on the basis of classical physics. In fact, given this planetary model we can even obtain a qualitative picture of the classical radiation spectrum. This also turns out to be wrong when compared with experiment.

From classical electromagnetic theory we find that a charge of magnitude e (in Coulombs) undergoing an acceleration a radiates energy at the rate

$$S = \frac{2}{3}\frac{e^2a^2}{c^3}\left[\frac{1}{4\pi\epsilon_0}\right] \tag{1.3.17}$$

where a is measured in m/s^2 and c, the speed of light, is measured in m/s. The charge of an electron in Coulombs is $e = 1.602191710 \times 10^{-19}$ C. If we use e.s.u. (electrostatic units) then the factor $[\frac{1}{4\pi\epsilon_0}]$ must be dropped. Later we shall use e.s.u. exclusively but in this chapter we give the formulas in MKS rationalized units with a square [ ] bracket for the factor to be dropped to get the corresponding formula in electrostatic units.

Now consider an hydrogen atom consisting of an electron in a spherical orbit about a proton. To a good approximation the centre of mass is located at the centre of the proton. The acceleration is given by:

$$a = \frac{v^2}{R} \tag{1.3.18}$$

where v is the speed of the electron and R the radius of its orbit.

Equating the mass times the centripetal acceleration and force of electrostatic attraction we get:

$$m\frac{v^2}{R} = \frac{e^2}{R^2}\left[\frac{1}{4\pi\epsilon_0}\right] . \tag{1.3.19}$$

Combining (1.3.18) and (1.3.19) yields

$$a = \frac{e^2}{mR^2}[\frac{1}{4\pi\epsilon_0}] \tag{1.3.20}$$

so that

$$S = \frac{2}{3}\frac{e^6}{m^2c^3R^4}[\frac{1}{4\pi\epsilon_0}]^3 . \tag{1.3.21}$$

From this we can estimate the time t that it would take for an electron to lose all its kinetic energy and spiral into the proton according to

$$t \approx \frac{\text{kinetic energy}}{S} . \tag{1.3.22}$$

Thus,

$$t \approx \frac{e^2}{R^2}[\frac{1}{4\pi\epsilon_0}]\frac{3}{2}\frac{m^2c^3R^4}{e^6}[\frac{1}{4\pi\epsilon_0}]^3 \tag{1.3.23}$$

or

$$t \approx \frac{3}{4}\frac{m^2c^3R^4}{e^4}[\frac{1}{4\pi\epsilon_0}]^2 . \tag{1.3.24}$$

The mass of an electron in MKS units is about $m = 9.1 \times 10^{-31}$ kg and a good approximation for the hydrogen atom radius is about 10^{-10} m. Substituting all of the above numbers in equation (1.3.24) yields

$$t \approx 4 \times 10^{-10}\, s . \tag{1.3.25}$$

Compared even to only 1 year = 3.1×10^7 s this prediction is wrong by an incredible factor of 10^{17} . In fact millions of years or 10^{14} s are more reasonable estimates so that the classical prediction is wrong by at least a factor of 10^{24}. Clearly classical physics contradicts the stability of atoms.

There is a second difficulty with the classical result. It has to do with the radiation spectrum. The radiation frequency ν is, in fact, determined by the angular frequency ω of rotation of the electron in its orbit according to

$$\omega = 2\pi\nu . \tag{1.3.26}$$

Now using that the acceleration of the electron in its orbit is given by

$$a = \omega^2 R \tag{1.3.27}$$

and using (1.3.21) we get

$$S = [\frac{1}{4\pi\epsilon_0}]\frac{2}{3}\frac{e^2R^2}{c^3}(2\pi)^4\nu^4 \,. \tag{1.3.28}$$

So we conclude that the spectrum of the radiated energy is continuous. This also contradicts the experimental fact, namely that atomic spectra consist of discrete series of very sharp lines.

Such lines had been studied for half a century and were well classified by the turn of this century. After studying the spectra of many atoms, Rydberg and Ritz [1.2] independently discovered a very important result. They found that the discrete frequencies observed could be expressed more simply. In fact all frequencies could be described by

$$\nu_{nm} = A_n - A_m \quad n, m = 1, 2, 3, \ldots \tag{1.3.29}$$

that is, a difference of two terms. Thus, far fewer terms A_n than frequencies are required. This so-called *Rydberg-Ritz Combination Principle* provided an important clue in the development of quantum mechanics.

1.4 Photoelectric Effect

A schematic diagram of the experimental arrangement for studying the photoelectric effect is shown in figure 1.2. We shine light of a fixed frequency ν on a clear metal surface. The retarding voltage is then increased to a voltage V_c at which the current, as measured by the galvanometer, ceases. We now summarize some of the experimental results that are observed.

1. The frequency of the light ν must be greater than some critical frequency ν_0 (even for zero retarding voltage) in order that photo-electrons be emitted.
2. For light intensities as low as 10^{-10} watts/m the delay in time for the photo-current to reach a steady state is less than 10^{-9} s.
3. For a fixed frequency, the photo-current is proportional to the intensity of the light.
4. The energy of the photo-electrons increases linearly with the frequency of the light.

We now consider this experiment in purely classical terms and see that the experimental results are in violent disagreement with the conclusion we obtain. Classically, the time-averaged energy density of an electromagnetic wave is given by

$$u = [4\pi\epsilon_0]\frac{1}{8\pi}|\mathbf{E}|^2 \tag{1.4.30}$$

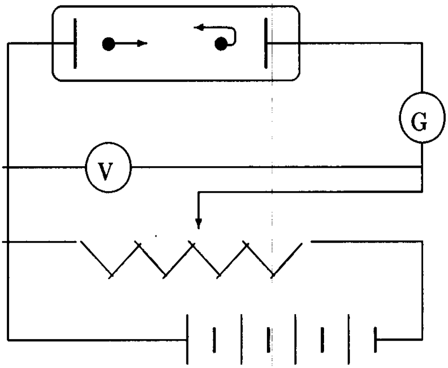

Figure 1.2: The photoelectric effect.

and is independent of frequency. Here **E** is the amplitude of the electric field in volts/m. Furthermore, classically, all electrons in the surface of the metal absorb energy uniformly. Thus, all the electrons eventually acquire enough energy to be emitted. It is therefore completely impossible to explain results 1. and 4. The experimental result 3. has some hope of being explained this way but it is not clear. The time delay, which is measured as 10^{-9} s or less for an intensity of 10^{-10} watt/m^2 can be estimated however. To do this we need a few numbers. The typical binding energy of an electron in the surface of a metal is 1 to 2 eV or about 10^{-19} Joules. The typical size of an atom in a metal is about 1Å$= 10^{-10}$ m. Thus, in 1m^2 of surface we have about 10^{20} atoms each of which requires an energy of 10^{-19} Joules. Thus about 10 Joules of energy must be absorbed by 1m^2 of metal surface before photo-electrons will be emitted. Since we are illuminating the surface with 10^{-10} watts/m^2 the time t required for photo-emission to occur is given by

$$10^{-10}t = 10 \tag{1.4.31}$$

or

$$1 \times t = 10^{11}\ s \approx 3{,}000 \text{ years.} \tag{1.4.32}$$

The discrepancy between theory and experiment is an incredible factor of 10^{20}.

Einstein [3] succeeded in explaining this discrepancy by making the same assumption as Planck, namely that light comes in quanta of energy given by

$$E = h\nu\ . \tag{1.4.33}$$

Thus, a given electron either absorbs all the energy E or none. Hence if ν is large enough so that E exceeds the binding energy ϕ of the electron, then photo-electrons will be emitted almost instantaneously. If, however, ν is too small, then no photo-electrons will be emitted at all. Thus the experimental results 1. and 4. are explained immediately. Result 3. also follows since the number of photons in the beam is what determines the number of photo-electrons emitted. But the number of photons N is given by

$$N = \frac{I}{h\nu} \tag{1.4.34}$$

where I is the intensity of the beam.

The final result 4. is also easily explained now. We simply equate the energy absorbed to the kinetic energy of the electron plus the energy required for the electron to break loose from the surface. This gives

$$\frac{1}{2}mv^2 + \phi = h\nu \tag{1.4.35}$$

which represents the linear relation between frequency and kinetic energy of the photo-electrons observed experimentally.

These results seem to indicate that photons are somehow particle-like, carrying a definite amount of energy, and are localized in space so that one atom can absorb a whole photon. In the subsequent section we discuss several more ideas and experiments that conclusively established the particle-like nature of photons.

1.5 Wave Particle Duality

Even in classical physics, light was not always considered as a wave motion. In fact, Newton formulated a completely corpuscular theory of light. The laws of reflection and refraction (Snell's Law) can then be derived purely on the basis of conservation of momentum and energy. The situation is as shown in figure 1.3. A particle with momentum p_1 is incident from the left in a homogeneous medium described by a constant potential V_1 and is either reflected or transmitted into a homogeneous medium described by a constant potential V_2. The particle thus experiences a force only at the interface between the media. This force is normal to the interface, since in a direction parallel to the interface $\mathrm{grad}V$ vanishes. Actually the particle receives an impulse on hitting the interface.

1.5.1 Reflection

Since the force is normal to the interface, the tangential component of momentum is conserved. Furthermore, since energy is conserved, the magnitude of the momentum $p_1 = \sqrt{2m(E - V_1)}$ is conserved. Hence, we get

$$p_1 \sin\theta_1 = p_1 \sin\theta_1' \ . \tag{1.5.36}$$

Thus,

$$\theta_1 = \theta_1' \tag{1.5.37}$$

so that the angle of incidence equals the angle of reflection.

1.5.2 Refraction

In this case we still have

$$E = \frac{{p_1}^2}{2m} + V_1 = \frac{{p_2}^2}{2m} + V_2 \ . \tag{1.5.38}$$

The tangential component of momentum is still conserved. Thus, we also have $p_1 = \sqrt{2m(E-V_1)}$ as before, and $p_2 = \sqrt{2m(E-V_2)}$. Equating the tangential components of momentum gives

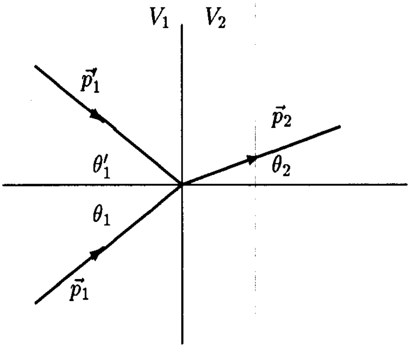

Figure 1.3: Snell's law.

$$p_1 \sin\theta_1 = p_2 \sin\theta_2 \tag{1.5.39}$$

so that,

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{p_2}{p_1} = \frac{\sqrt{2m(E-V_2)}}{\sqrt{2m(E-V_1)}} = \frac{n_2}{n_1} \ . \tag{1.5.40}$$

Here we have defined n the index of refraction to be $\sqrt{2m(E-V)}$. This is reasonable since $\sqrt{2m(E-V)}$ is simply a number for a given medium and hence also for a particle of a given energy.

If one uses wave optics one finds that the index of refraction of a medium is inversely proportional to the wavelength of the light in the medium. Thus,

$$\frac{n_2}{n_1} = \frac{\lambda_2}{\lambda_1} \ . \tag{1.5.41}$$

This suggests that for a particle of momentum p there may be an associated wavelength λ such that

$$\lambda \propto \frac{1}{p} \ . \tag{1.5.42}$$

1.6 de Broglie's Hypothesis

de Broglie presented a very simple argument for photons to fix the constant of proportionality in the above relationship. The relationship de Broglie used connects the energy E and momentum p for a plane electromagnetic wave. It states

$$E = cp \tag{1.6.43}$$

where c is the speed of light. This equation follows from the fact that photons travel at the speed of light c and have an effective momentum

$$p = \frac{E}{c^2}c = \frac{E}{c} \tag{1.6.44}$$

as stated in equation (1.6.43). If we combine this relationship with the Planck energy-frequency relation $E = h\nu$, we get:

$$h\nu = cp \ . \tag{1.6.45}$$

Thus,

$$\lambda = \frac{c}{\nu} = \frac{h}{p} \ . \tag{1.6.46}$$

This is the famous de Broglie hypothesis, namely that with any particle carrying momentum p one should associate a wave-length $\lambda = h/p$.

1.7 The Compton Effect

The photoelectric effect indicated that photons were somehow granular or particle-like carrying a definite amount of energy given by $E = h\nu$. By scattering X-rays (photons) off free electrons, A.H. Compton [4] showed that photons are definitely particle-like, carrying a definite momentum and scattering like point particles. The situation is as depicted in figure 1.4.

The incoming photon carries energy $E = h\nu$ and momentum $p = (h\nu)/c$ in the x-direction, while the electron is initially at rest (energy $E = mc^2$). The scattered photon carries energy $E' = h\nu'$ and total momentum $p' = (h\nu')/c$ in the direction given by the angle θ. The electron, which was originally at rest, recoils with momentum p in the direction given by the angle φ. Applying conservation of energy and momentum as for point particles we get:
Conservation of Energy

$$h\nu + mc^2 = h\nu' + E \ . \tag{1.7.47}$$

Conservation of Momentum

$$\frac{h\nu}{c} = \frac{h\nu'}{c}\cos\theta + p\cos\varphi \tag{1.7.48}$$

$$0 = \frac{h\nu'}{c}\sin\theta - p\sin\varphi \ . \tag{1.7.49}$$

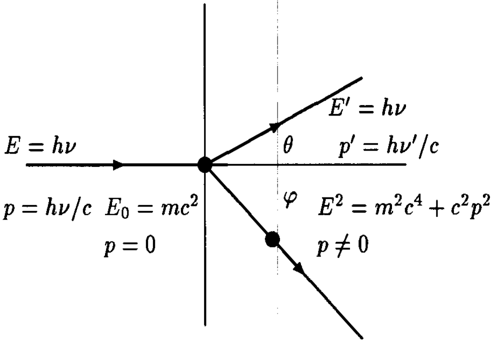

Figure 1.4: The Compton effect.

In his experiments Compton measured the change in wavelength of the scattered X-rays as a function of the scattering angle θ. According to classical electromagnetic theory no change in wavelength should occur. Now in writing down equations (1.7.47) - (1.7.49) we have treated the X-rays as if they were point particles. Before further discussion we first obtain a formula for the change in wavelength $\Delta\lambda = \lambda - \lambda'$ in terms of θ. By rearranging (1.7.48) and (1.7.49), squaring and adding we get:

$$h^2\nu^2 + h^2\nu'^2 - 2h^2\nu\nu'\cos\theta = c^2p^2 \tag{1.7.50}$$

Rearranging (1.7.47) and squaring yields:

$$h^2\nu^2 + h^2\nu'^2 - 2h^2\nu\nu'\cos\theta = E^2 + m^2c^4 - 2Emc^2 \tag{1.7.51}$$

Subtracting (1.7.50) from (1.7.51) and using the energy momentum relation

$$E^2 = c^2p^2 + m^2c^4 \tag{1.7.52}$$

we get:

$$-2h^2\nu\nu'(1 - \cos\theta) = -2mc^2(E - mc^2) . \tag{1.7.53}$$

But from equation (1.7.47) we have

$$E - mc^2 = h(\nu - \nu') . \tag{1.7.54}$$

Thus, we finally get:

$$\frac{h}{mc}(1 - \cos\theta) = c\frac{\nu - \nu'}{\nu\nu'} = \frac{c}{\nu'} - \frac{c}{\nu} = \lambda - \lambda' . \tag{1.7.55}$$

So the increase in wavelength is given by

$$\Delta\lambda = \lambda' - \lambda = \lambda_C(1 - \cos\theta) . \tag{1.7.56}$$

Here we have introduced the *Compton wavelength* of the electron, given by

$$\lambda_C = \frac{h}{mc} \,. \tag{1.7.57}$$

Compton's measurements showed that equation (1.7.56) agreed splendidly with the experimental results. Thus, a photon has particle properties just like a particle has wave properties. In the next section we discuss an experiment (the Davisson-Germer experiment) that proved conclusively the wave-like nature of particles. So it was found that on the one hand waves sometimes behaved like particles whereas on the other hand particles sometimes behaved like waves.

1.8 The Davisson-Germer Experiment

If we consider the de Broglie relation

$$\lambda = \frac{h}{mv} \tag{1.8.58}$$

for an electron with 100 eV energy then we find

$$100\,eV = 100 \times 1.6 \times 10^{-19}\,J = \frac{1}{2}mv^2 \tag{1.8.59}$$

so with $m = 9.11 \times 10^{-31}$ kg we get

$$mv = 5.40 \times 10^{-24}\ \text{kg m/s} \tag{1.8.60}$$

and

$$\lambda = \frac{h}{mv} = 1.23 \times 10^{-10}\ \text{m/s} \tag{1.8.61}$$

or

$$\lambda = 1.23\,\text{Å} \,. \tag{1.8.62}$$

Thus, λ is about the same as the wavelength of a hard X-ray.

Now the diffraction of X-rays by crystals had already been observed and explained in 1913 by the father and son team of W.L. and W.H. Bragg [1.5]. The planes of a crystal lattice act as a very fine diffraction grating. Secondly, de Broglie during his thesis defence suggested that the matter waves predicted by his formula could be observed in such a manner. The actual experiment, however, was carried out almost simultaneously by Davisson and Germer as well as G.P. Thomson [6]. Davisson and Germer produced such a diffraction pattern with an electron beam without knowing about the de Broglie hypothesis. (This experiment was also performed by E. Rupp [7] using a ruled grating and grazing incidence.) Davisson and Germer had been scattering electrons off polycrystalline nickel targets when a fortunate accident occurred. The vacuum system broke down and their nickel target oxidized. After repairing the vacuum they tried to expel the oxygen from the nickel by heating. This process changed the polycrystalline nickel to several large crystals. On recommencing the scattering experiment they found that 54 volt electrons incident at 50° to the surface of the nickel, led to extremely strong reflection of these electrons. The result had all the appearance of Bragg reflections. Now from X-ray data the lattice spacing $2a$ for nickel was known to be 2.15 Å$= 2.15 \times 10^{-10}$ m. The Bragg formula (derived below) states that for reflection maxima the incident angle should be given by:

$$n\lambda = 2a\sin\theta \qquad n = 1, 2, 3, ... \tag{1.8.63}$$

Using $n = 1,\ 2a = 2.15\text{Å},\ \theta = 50°$ yields:

$$\lambda = 1.65\,\text{Å} \tag{1.8.64}$$

Computing the wavelength of a 54 volt electron from

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2meV}} \tag{1.8.65}$$

we find

$$\lambda = 1.67\,\text{Å} \tag{1.8.66}$$

This is splendid confirmation of the de Broglie hypothesis. We now derive the Bragg formula . The situation is as depicted in figure 1.5. The small spheres represent individual atoms. We begin by computing the path difference Δl for the two rays reflected as shown. This path difference is given by:

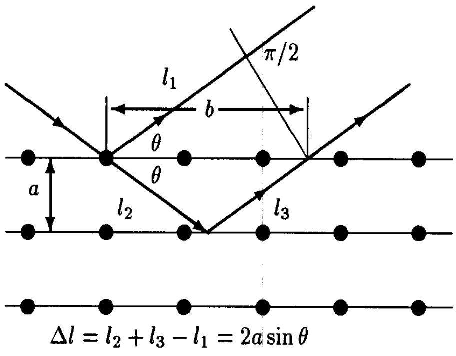

Figure 1.5: Bragg reflection.

$$\Delta l = l_2 + l_3 - l_1 \tag{1.8.67}$$

Now,

$$l_2 = l_3 = \frac{a}{\sin\theta} \tag{1.8.68}$$

Also,

$$b = l_2 \cos\theta + l_2 \cos\theta = 2a \cot\theta \ . \tag{1.8.69}$$

But,

$$l_1 = b \cos\theta = 2a \frac{cos^2\theta}{\sin\theta} \ . \tag{1.8.70}$$

Therefore,

$$\Delta l = \frac{2a}{\sin\theta}(1 - \cos^2\theta) = 2a \sin\theta \ . \tag{1.8.71}$$

For constructive interference (maxima) we need that

$$\Delta l = n\lambda \qquad n = 1, 2, 3, \ldots \ . \tag{1.8.72}$$

Thus, we have arrived at the Bragg formula.

1.9 The Franck-Hertz Effect

This effect was first observed in an experiment in 1914 in which J. Franck and G. Hertz [8] looked for the "grainyness" of matter suggested by Planck's radiation law. A later experiment in 1924 by G. Hertz [9] verified Bohr's atomic model predictions. We describe both experiments at the same time since the second experiment was simply a more detailed investigation of the first. A schematic diagram of the experimental set-up is shown in figure 1.6. Electrons are emitted from the hot cathode c and accelerated by the potential difference V between the cathode c and the grid g towards the anode a. The current I reaching the anode a is measured by the ammeter A. The current observed as a function of the accelerating voltage is displayed in figure 1.7. The first drop in current occurs at 4.9 eV, the second at 9.8 eV and the third at 14.7 eV. Thus, it appears that the mercury atoms absorb 1×4.9 or 2×4.9 or 3×4.9 eV of energy from the accelerated electrons. Now the most prominent spectral line emitted by mercury is in the ultraviolet at $\lambda = 2537$ Å.

If we use the Planck energy-frequency relation $E = h\nu = (hc)/\lambda$ we find that this wavelength corresponds to an energy

$$E = \frac{6.63 \times 10^{-34} \times 3.00 \times 10^8}{2.54 \times 10^{-7}} = 7.83 \times 10^{-19}\ \mathrm{J} = 4.9\ \mathrm{eV} \ . \tag{1.9.73}$$

This result can be interpreted as follows. For an energy below 4.9 eV the electrons make essentially elastic collisions with the mercury atoms and the current increases with voltage. At 4.9 eV the electrons can collide inelastically with the mercury atoms to excite the 2537 Å line and thus give up all their kinetic energy. Thus, the current drops. Then as the voltage increases further the current again rises until we reach a voltage of $9.8 = 2 \times 4.9$ volts. At this point the electrons can make two inelastic collisions and lose all their kinetic energy. The process again repeats until at $14.7 = 3 \times 4.9$ volts the electrons can make

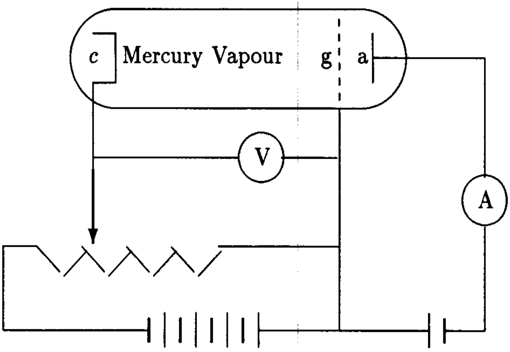

Figure 1.6: The Franck-Hertz Experiment.

three inelastic collisions. To verify this interpretation G. Hertz [9] measured the radiation emitted by the mercury atoms and found that below 4.9 volts the ultraviolet radiation corresponding to $\lambda = 2537$ Å was not emitted, whereas above 4.9 volts it was emitted strongly. In fact, Hertz was able to correlate the occurrence of many of the spectral lines of mercury with a threshold voltage of the electrons. Thus, the quantized nature of the energy levels of an atom was definitely established.

1.10 Planck's Radiation Law

We now return to a discussion of blackbody radiation. Planck combined classical statistical mechanics with his energy frequency relation. Thus he made the following assumptions:

1. In accordance with classical statistical mechanics, the probability for an oscillator of energy E to be excited is proportional to the Boltzmann factor $\exp(-E/k_BT)$.
2. The energy of the oscillators is quantized and comes in quanta given by $\langle E\rangle = nh\nu$, $n = 1, 2, 3, \ldots$.

Combining these assumptions we can compute the average energy $\langle E\rangle$ of an oscillator from
$\langle E\rangle$ = Sum [(Probability of energy E_i)× (energy E_i)]
so

$$\langle E\rangle = \frac{\sum_n nh\nu e^{-nh\nu/k_BT}}{\sum_n e^{-nh\nu/k_BT}} \ . \tag{1.10.74}$$

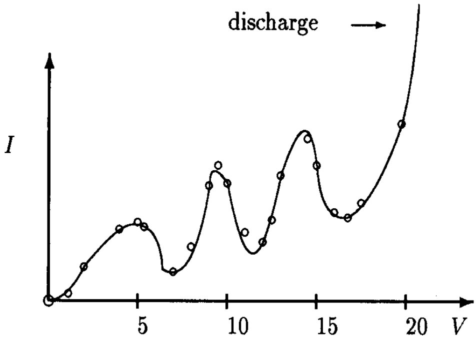

Figure 1.7: Current-Voltage Curve for the Franck-Hertz Experiment: The ordinate is the current (in arbitrary units) as measured by the ammeter A.

This can be written

$$\langle E\rangle = -\frac{d}{d(\frac{1}{k_BT})}\ln\sum_n e^{-nh\nu/k_BT} \ . \tag{1.10.75}$$

After carrying out the sum of the geometric series we get

$$\langle E\rangle = -\frac{d}{d(\frac{1}{k_BT})}\ln\,[1-e^{-h\nu/k_BT}]^{-1} \ . \tag{1.10.76}$$

Thus,

$$\langle E\rangle = k_BT\frac{h\nu/k_BT}{e^{h\nu/k_BT}-1} \ . \tag{1.10.77}$$

We next combine this result with the result obtained in section 1.1 and write the equation.

> The energy per unit volume du in a frequency interval between ν and $\nu + d\nu$
> $= \langle E\rangle \times$ (number of modes of oscillation in this interval) .

After dividing by the volume V of the blackbody cavity we find from equation (1.2.12) that

$$\frac{1}{V}dN = \frac{8\pi}{c^3}\nu^2 d\nu \ . \tag{1.10.78}$$

Thus,

$$du = \langle E\rangle\frac{1}{V}dN = \frac{8\pi}{c^3}\frac{h\nu^3}{e^{h\nu/(k_BT)}-1}d\nu \tag{1.10.79}$$

so

$$\rho(\nu) = \frac{8\pi}{c^3}\frac{h\nu^3}{e^{h\nu/(k_BT)} - 1} \ . \tag{1.10.80}$$

This is the famous Planck's blackbody radiation formula shown in fig. (1.1). We notice that for low frequencies namely $h\nu/(k_BT) << 1$, this law just goes over into the Rayleigh-Jeans law since for these frequencies we have $\exp(h\nu/(k_BT)) \approx 1 + h\nu/(k_BT)$.

The establishment of this formula and the introduction of the constant h was one of the most revolutionary developments in all of physics.

1.11 Einstein's Model for Specific Heat

Soon after Planck's enunciation of his blackbody law, Einstein [1.10] in 1907 used a very similar argument to explain why the specific heat of solids goes to zero at zero temperature. Einstein's model was only qualitatively correct, a better model was published in 1912 by P. Debye [1.11]. This so-called Debye model remains valid today. We shall briefly discuss this model after considering Einstein's model. From the classical equipartition theorem one obtains for the internal energy U of one mole of a monatomic solid at temperature T the value

$$U = \frac{1}{2}k_BT \times \text{(number of degrees of freedom)} \tag{1.11.81}$$

or

$$U = 3N_Ak_BT \tag{1.11.82}$$

where N_A is Avagadro's number 6.02×10^{23} /mole. Here we have assumed that the atoms are harmonically bound so that associated with each degree of freedom for the kinetic energy there is also a degree of freedom for the potential energy. That is why we again have k_BT without the factor of 1/2. Now, $N_Ak_B = R$ is the molar gas constant with a value of ≈ 2 cal/mole.K . Hence, the specific heat at constant volume C_V is given by

$$C_V = (\frac{\partial U}{\partial T})_V = 3R \approx 6 \quad \text{cal/mole.K} \ . \tag{1.11.83}$$

This result, known as the law of Dulong and Petit, agrees very well at high temperature but fails miserably for low temperatures where C_V goes to zero like T^3.

By assuming that the atoms in a monatomic crystal are harmonically bound (simple oscillators) and applying Planck's assumptions to these oscillators Einstein obtained that the average energy $\langle E\rangle$ of such an oscillator is given by

$$\langle E\rangle = k_BT\frac{h\nu/(k_BT)}{e^{h\nu/(k_BT)} - 1} \ . \tag{1.11.84}$$

Now applying the equipartition principle and using the fact that for one mole of a monatomic solid the number of degrees of freedom is given by $3N_A$ he found that the total internal energy U is given by

$$U = 3N_A k_B T \frac{h\nu/(k_B T)}{e^{h\nu/(k_B T)} - 1} \tag{1.11.85}$$

Clearly at high temperatures $(h\nu/(k_B T) < 1)$ and we again have

$$e^{h\nu/(k_B T)} \approx 1 + h\nu/(k_B T)$$

and thus,

$$U \approx 3N_A k_B T = 3RT \tag{1.11.86}$$

in accord with experiment and the classical result. At low temperatures $(h\nu/(k_B T) > 1)$ we have

$$e^{h\nu/(k_B T)} - 1 \approx e h\nu/(k_B T)$$

and thus,

$$U \approx 3N_A h\nu e^{-h\nu/(k_B T)} . \tag{1.11.87}$$

Thus, both U and $C_V = (\partial U/\partial T)_V$ vanish exponentially at low temperature. The agreement with experiment is certainly much better than the classical result.

1.12 The Debye Model

The reason that U vanishes too rapidly at low T was pointed out by P. Debye [1.11]. He assumed that the vibration spectrum of a monatomic solid may be treated, as we have done above, as a homogeneous medium except that the total number of modes is cut off at $3N$, where N is the number of atoms per unit volume ($= N_A$ for one mole), to yield the correct number of degrees of freedom of N atoms. This yields a cutoff frequency ν_{max}. Now the number of modes per unit volume with frequency less than ν_{max} is given by a formula like (1.2.11) namely,

$$3N = \frac{1}{8} \times \frac{4}{3}\pi \left(\frac{2\nu}{v_0}\right)^3 \times 3 . \tag{1.12.88}$$

Here v_0 is an *average speed of sound* in the solid and the last factor of 3 is due to the fact that there are two transverse modes plus one longitudinal mode of vibration in a solid. We have replaced the two different velocities for these modes by one "average" velocity. Thus, the cutoff frequency is given by

$$\nu_{max} = v_0 \left(\frac{3N_A}{4\pi}\right)^{1/3} \tag{1.12.89}$$

The internal energy per unit volume, dU, for frequencies between ν and $\nu + d\nu$ is now given by the number of modes in this interval times the average energy per mode, namely

$$dU = \frac{12\pi}{v_0^3}\nu^2 d\nu \frac{h\nu}{e^{h\nu/(k_BT)} - 1} \tag{1.12.90}$$

Thus, the total internal energy is given by

$$U = \frac{12\pi}{v_0^3}\int_0^{\nu_{max}} \frac{h\nu^3 d\nu}{e^{h\nu/(k_BT)} - 1} \tag{1.12.91}$$

Letting,

$$x = \frac{h\nu}{k_BT} \, , \; x_{max} = \frac{h\nu_{max}}{k_BT} \tag{1.12.92}$$

we find

$$U = \frac{12\pi(k_BT)^4}{h^3{v_0}^3}\int_0^{x_{max}} \frac{x^3 dx}{e^x - 1} \, . \tag{1.12.93}$$

This is Debye's expression. Notice that for very low temperatures $x_{max} \to \infty$ and so the integral just becomes a number which can be calculated and turns out to be about 2.40437. Thus, U is proportional to T^4 and therefore, C_V is proportional to T^3 for low temperatures. The dimensionless parameter x_{max} can be rewritten as

$$x_{max} = \frac{\theta_D}{T} \tag{1.12.94}$$

where the *Debye temperature* θ_D is given by

$$\theta_D = \frac{hv_0}{k_B}\left(\frac{3N_A}{4\pi}\right)^{1/3} \tag{1.12.95}$$

The Debye temperature characterizes the specific crystal through the velocity v_0. Typical values of θ_D range from 150 K to 1000 K.

1.13 Bohr Model and the Hydrogen Atom

In this section we describe a model that was a precursor of quantum mechanics. Bohr extended Planck's hypothesis and made some additional assumptions. We describe these now and apply them to the hydrogen atom in the next section.

1. To get the observed stability of atoms, Bohr assumed that atoms exist only in certain definite states in which they do not radiate. These are the *stationary states.* The energy is therefore automatically quantized since only these stationary states occur and not all possible states.

2. To get discrete spectra Bohr assumed Planck's law in the form

$$E_j - E_k = h\nu \, . \tag{1.13.96}$$

Here E_j, E_k are the energies associated with two stationary states. Thus, he made the implicit further assumption that the energy changes discontinuously from one state to another. This explains discrete spectra both for emission and absorption:

Emission occurs if $E_j > E_k$
Absorption occurs if $E_k > E_j$.

This also gives the Rydberg-Ritz combination principle immediately since from the combination principle

$$\nu = A_j - A_k \tag{1.13.97}$$

Thus, the terms A_j can be identified with E_j/h .

3. The correspondence principle was one of Bohr's most useful assumptions. It states: In the limit of large quantum numbers the classical predictions must be recovered at least asymptotically.

A motivation for this assumption can be found in the experimental fact that

$$A_n = \frac{a}{(n+b)^2} \tag{1.13.98}$$

for almost all atomic spectra. Thus,

$$E_n = \frac{ah}{(n+b)^2} \quad . \tag{1.13.99}$$

Then,

$$E_n - E_m = \frac{ah}{(n+b)^2} - \frac{ah}{(m+b)^2} \tag{1.13.100}$$

and

$$E_n - E_m \underset{n,m\to\infty}{\longrightarrow} ah\frac{m^2-n^2}{m^2n^2} \longrightarrow 0 \; . \tag{1.13.101}$$

Hence, $\Delta E_n \longrightarrow 0$ and we get a continuum of energies for large quantum numbers.

We now show that the correspondence principle leads to quantization of orbital angular momentum and then use this result to derive the hydrogen spectrum.

We assume the electron in an hydrogen atom is in a circular orbit about a force centre given by the Coulomb attraction of the proton. Also we choose the zero of energy to correspond to an unbound electron with zero kinetic energy. Thus, in the bound orbit E is negative. Classically the frequency ω at which the electron radiates is given by the angular frequency of rotation of the electron in its orbit. Thus,

$$\omega = \frac{v}{r} \; . \tag{1.13.102}$$

The total energy of the electron in its orbit is

$$E = \frac{1}{2}mv^2 - [\frac{1}{4\pi\epsilon_0}]\frac{e^2}{r} \; . \tag{1.13.103}$$

Equating the Coulomb force of attraction to the mass times the centripetal acceleration gives

$$m\frac{v^2}{r} = [\frac{1}{4\pi\epsilon_0}]\frac{e^2}{r^2} \; . \tag{1.13.104}$$

Thus,

$$E = -\frac{1}{2}mv^2 = -[\frac{1}{4\pi\epsilon_0}]\frac{e^2}{2r} \; . \tag{1.13.105}$$

Also,

$$L = mvr = m\omega r^2 \tag{1.13.106}$$

and since

$$m\frac{v^2}{r} = [\frac{1}{4\pi\epsilon_0}]\frac{e^2}{r^2} \tag{1.13.107}$$

we find that

$$mv^2r = [\frac{1}{4\pi\epsilon_0}]e^2 \quad \text{or} \quad m^2v^2r^2 = L^2 = [\frac{1}{4\pi\epsilon_0}]mre^2 \; . \tag{1.13.108}$$

Therefore,

$$r = [4\pi\epsilon_0]\frac{L^2}{me^2} \; . \tag{1.13.109}$$

Then,

$$E = -\frac{1}{2}\left[\frac{1}{4\pi\epsilon_0}\right]^2 \frac{me^4}{L^2} \tag{1.13.110}$$

and

$$\frac{dE}{dL} = \left[\frac{1}{4\pi\epsilon_0}\right]^2 \frac{me^4}{L^3} \; . \tag{1.13.111}$$

But, as we found above,

$$L = m\omega r^2 = m\omega\left([4\pi\epsilon_0]\frac{L^2}{me^2}\right)^2 = \omega\frac{[4\pi\epsilon_0]^2L^4}{me^4} \; . \tag{1.13.112}$$

So,

$$\left[\frac{1}{4\pi\epsilon_0}\right]^2 \frac{me^4}{L^3} = \omega \ . \tag{1.13.113}$$

This means that the change in E when L changes is given by

$$\frac{\Delta E}{\Delta L} = \omega \ . \tag{1.13.114}$$

Now, the smallest change that is possible in E according to Planck is $\Delta E = h\nu = \hbar\omega$. Therefore, we get that

$$\Delta E = h\nu = 2\pi\nu\Delta L \ . \tag{1.13.115}$$

This in turn means that the smallest change possible in L is given by

$$\Delta L = \frac{h}{2\pi} = \hbar \, . \tag{1.13.116}$$

So, the angular momentum is quantized and we have

$$L = n\hbar \, . \tag{1.13.117}$$

Replacing this back in the expression (1.13.110) for the energy E we get

$$E(n) = -\frac{1}{2}\left[\frac{1}{4\pi\epsilon_0}\right]\frac{me^4}{n^2\hbar^2} \, . \tag{1.13.118}$$

Thus, we have calculated the terms in the Ritz combination principle for hydrogen. The result agrees with experiment to better than 0.01% . As a side benefit we also get

$$r_n = [4\pi\epsilon_0]\,\frac{n^2\hbar^2}{me^2}n^2 = a_0\, n^2 \, . \tag{1.13.119}$$

Here we have introduced the Bohr radius

$$a_0 = [4\pi\epsilon_0]\,\frac{n^2\hbar^2}{me^2} = 0.0529 \ \text{nm} \, . \tag{1.13.120}$$

This completes a brief review of the major developments that led to the modern quantum theory. In the next chapter we briefly review certain aspects of classical mechanics that turned out to be particularly useful in developing an understanding of quantum theory.

Up to this point we have used MKS rationalized electromagnetic units. In atomic physics most calculations are performed in Gaussian units. Thus, charge is measured in esu. so that $e \approx 4.8 \times 10^{-10}$ esu. and magnetic fields are measured in Gauss. To convert our previous equations to Gaussian units simply drop the square factors $[4\pi\epsilon_0]$ or $\frac{1}{[4\pi\epsilon_0]}$ appearing in equations involving electromagnetism.

1.14 Problems

1.1 Calculate the principle quantum number for the earth in its orbit about the sun. What is the energy difference between two neighbouring energy levels? Hint: For large $n, E_n \approx E_{classical}$.

1.2 What is the wavelength associated with gas molecules at a temperature T ? Estimate this wavelength for a typical gas at room temperature and compare it to visible light.

1.3 For a monochromatic beam of electromagnetic radiation ($\lambda \approx 5000$ Å) of an intensity of 1watt/m^2, calculate the number of photons passing 1 cm^2 of area normal to the beam in one second.

1.4 Show that if one assumes that the circumference of a stationary state orbit of an electron in an hydrogen atom is an integral multiple of the de Broglie wavelength, one also obtains the correct energy levels equation (1.13.115).

1.5 List several experiments or observational results that may be used to obtain a lower bound on the lifetime of atomic hydrogen in free space.

1.6 Estimate the effect on the specific heat of reducing a crystal to a fine powder of dimensions of about 10^{-6} cm.
Hint: Study the Debye model of specific heat and realize that the size of the crystal now also imposes an upper limit on the wavelength of the sound waves in the crystal.

1.7 The shortest possible wavelength of sound in sodium chloride is twice the lattice spacing, about 5.8×10^{-8} cm. The sound velocity is approximately 1.5×10^5 cm/sec.
a) Compute a rough value for the highest sound frequency in the solid.
b) Compute the energy of the corresponding phonons, or quanta of vibrational energy.
c) Roughly what temperature is required to excite these oscillations appreciably?

Bibliography

[1.1] A. B. Pippard - *The Elements of Classical Thermodynamics* - Cambridge University Press (1957) page 78.

[1.2] J.R. Rydberg, Report of Intl. Phys. Cong. at Paris, ii, 200 (1900). W. Ritz, Phys. Z. **9**, 521 (1908). W. Ritz, Astrophys. J. **28**, 237 (1908).

[1.3] A. Einstein, Ann. Phys. **17**, 132 (1905).

[1.4] A.H. Compton, Phys. Rev. **22**, 409 (1923).

[1.5] W.H. Bragg and W.L. Bragg, Proc. Roy. Soc. **A88**, 428 (1913). W.L. Bragg, Proc. Roy. Soc. **A89**, 248 (1913).

[1.6] C. Davisson and L.H. Germer, Phys. Rev. **30**, 705 (1927).
G.P. Thomson, Proc. Roy. Soc. **A117**, 600 (1928).

[1.7] E. Rupp, Zeits. f. Phys. **52**, 8 (1928), also Phys. Zeits. **28**, 837 (1928)

[1.8] J. Franck and G. Hertz, Verh. Dtsch. Phys. Ges. **16**, 512 (1914).

[1.9] G. Hertz, Z. Phys. **22**, 18 (1924).

[1.10] A. Einstein, Ann. d. Physik **22**, 180 (1907); **34**, 170 (1911).

[1.11] P. Debye, Ann. d. Physik **29**, 789 (1912).

For further reading on this subject we strongly recommend the following books:

A.P. French and E.F. Taylor - *An Introduction to Quantum Physics* - W.W. Norton and Co. Inc. (1978).

A.T. Weidner and R.L. Sells - *Elementary Modern Physics* - Allyn and Bacon, 2nd edition (1969).

M. Born - *Atomic Physics* - Blackie and Son Ltd., London, 5th edition (1952).

C. Kittel - *Introduction to Solid State Physics* - John Wiley and Sons, 2nd edition (1963) Chapter 6.

Chapter 2

Review of Classical Mechanics

2.1 Introduction

Classical mechanics was reformulated in an elegant manner during the nineteenth century. In this chapter we review some of these formulations. To lead into the discussion we consider in section 2.2 the mechanics of a point particle confined to one dimension and acted on by a conservative force. We use this to give an heuristic introduction to Hamilton's principle and the Euler-Lagrange equations as well as Hamilton's equations. In section 2.3, we review the Lagrangian and Hamiltonian formulation of many-particle systems. This section assumes a somewhat deeper knowledge of analytical mechanics. The discussion of section 2.3 is further elaborated to Hamilton- Jacobi theory in sections 2.4 and 2.5. In section 2.6, we then apply some of these classical techniques to find constants of the motion and perform Bohr-Sommerfeld quantization for the hydrogen atom. Finally we give a heuristic derivation, actually more of a plausibility argument, for the Schrödinger equation in section 2.7.

2.2 Classical Mechanics: Particle in One Dimension

We begin by considering a particle of mass m moving in one dimension and acted on by a conservative force $F(x)$. The equation of motion is given by:

$$m\frac{d^2x}{dt^2} = F(x) \tag{2.2.1}$$

This equation contains all the information about the dynamics (forces) of this system. In addition (since equation (2.2.1) is second order in time) we need two pieces of initial data, say the position x_0 and velocity v_0 of the particle at some instant t_0, in order to determine the motion of the system (particle) for all subsequent times. We now restate this simple fact in somewhat more fancy language which will later allow us to compare classical and quantum mechanics more easily.

To begin with we define $x(t)$ and $v(t) = dx/dt$ to be dynamical variables. Because equation (2.2.1) allows us to express $a(t) = d^2x/dt^2$ in terms of $x(t)$, it follows that all dynamical variables such as acceleration, energy, momentum, etc. may be expressed in terms of $x(t)$ and $v(t)$. That we need both $x(t)$, and $v(t)$ follows from the fact that (2.2.1) is second order in time. We therefore say that $x(t)$ and $v(t)$ form a complete set of dynamical variables. Specifying a complete set of dynamical variables (i.e. $x(t)$ and $v(t)$) at a given time t_0, specifies the state of the system completely at the instant t_0. This means that if we know $x(t_0)$ and $v(t_0)$ then all physical quantities at $t = t_0$ are, in principle, determined.

To determine the state of the system at any later time $t > t_0$ we need to solve equation (2.2.1) for $t > t_0$. We want to bring out more clearly this fact that x and v specify the state completely. To this end we want to find a pair of first order differential equations for a complete set of dynamical variables. It turns out that rather than x and v it is more convenient to use x and p. Here p is the linear momentum. [1]

$$p = mv \ . \tag{2.2.2}$$

We now rewrite equation (2.2.1) using the fact that the force F is conservative. This means that there is a potential $V(x)$ such that

$$F = -\frac{\partial V}{\partial x} \ . \tag{2.2.3}$$

Then, (2.2.1) reads:

$$m\frac{d^2x}{dt^2} = -\frac{\partial V}{\partial x} \ . \tag{2.2.4}$$

This equation may be immediately integrated if we multiply by $\frac{dx}{dt}$. We then get

$$\frac{1}{2}m\frac{d}{dt}\left(\frac{dx}{dt}\right)^2 = -\frac{\partial V}{\partial x}\cdot\frac{dx}{dt} \tag{2.2.5}$$

so that integrating between t_0 and t we find

$$\frac{1}{2}mv^2(t) + V(x) = \frac{1}{2}mv^2(t_0) + V(x_0) = E \ . \tag{2.2.6}$$

---

[1] Equation (2.2.2) defines the momentum correctly for this example. In general, the "canonical" momentum is involved and a more general definition is required. This is done in detail in the next section, equation (2.3.39). See also equation (2.2.31).

Here E is a constant of integration. It represents the total energy. If we solve equation (2.2.2) for v in terms of p and substitute the result in equation (2.2.6) we get:

$$\frac{p^2}{2m} + V(x) = E \ . \tag{2.2.7}$$

The function of p and x on the left hand side of this equation still represents the total energy but as a function of p and x it is called the Hamiltonian $H(x,p)$ of this system.

$$H(x,p) = \frac{p^2}{2m} + V(x) \ . \tag{2.2.8}$$

If we know the Hamiltonian, it is a simple matter to find a pair of equations equivalent to the definition (2.2.2) and equation of motion (2.2.4). In fact by straightforward differentiation we see that:

$$\frac{\partial H}{\partial p} = \frac{p}{m} = \frac{dx}{dt} \tag{2.2.9}$$

and

$$\frac{\partial H}{\partial x} = \frac{\partial V}{\partial x} = -F = -\frac{dp}{dt} \ . \tag{2.2.10}$$

The pair of equations

$$\frac{dx}{dt} = \frac{\partial H}{\partial p} \tag{2.2.11}$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x} \tag{2.2.12}$$

are called Hamilton's equations of motion. At this stage they appear to be nothing more than a fancy way of rewriting (2.2.1) and (2.2.2). In 1834 and 1835, Sir William Rowan Hamilton [2.1] published two papers which gave deep insight into classical mechanics. We give a heuristic discussion of this Hamilton's principle before we state it.

Consider a simple harmonic oscillator (consisting of a point mass m and spring with constant k), which starts from rest with the spring compressed. Thus, all the energy is initially potential and none of the energy is kinetic.

$$T(0) = 0 \ , \quad V(0) = E \tag{2.2.13}$$

After 1/4 of a cycle τ the spring is completely unstressed and the potential energy is zero. On the other hand the particle is moving with maximum velocity so that the kinetic energy is a maximum, in fact the total energy:

$$T(\frac{\tau}{4}) = E \ , \quad V(\frac{\tau}{4}) = 0 \ . \tag{2.2.14}$$

After 1/2 cycle the situation is again reversed. The particle has overshot the equilibrium and the spring is stretched with the particle momentarily at rest so that

$$T(\frac{\tau}{2}) = 0 \ , \ V(\frac{\tau}{2}) = E \ . \tag{2.2.15}$$

After 3/4 cycle we have:

$$T(\frac{3\tau}{4}) = E \ , \ V(\frac{3\tau}{4}) = 0 \ , \tag{2.2.16}$$

and after a full cycle we are back to our initial configuration

$$T(\tau) = 0 \ , \ V(\tau) = E \ . \tag{2.2.17}$$

If we consider this motion, we see that as it progresses, we alternate between potential energy V and kinetic energy T. Thus, the motion is such that the function

$$L = \int_{t_1}^{t_2} (T - V) dt \tag{2.2.18}$$

of the particle path is minimized. We call L the action and

$$\mathcal{L} = T - V \tag{2.2.19}$$

the Lagrangian of the system. Hamilton's principle states:
A dynamical system evolves along that path which minimizes or maximizes the action. To see how this works let us consider our one-dimensional particle.

$$L = \int_{t_1}^{t_2} \left[\frac{1}{2} m \dot{x}^2 - V(x) \right] dt \tag{2.2.20}$$

where we are writing $\dot{x}$ for dx/dt . In extremizing (maximizing or minimizing) L we must choose $\dot{x}(t), x(t)$ such as to pass through $\dot{x}(t_1), x(t_1)$ and $\dot{x}(t_2), x(t_2)$. The state at the end points t_1 and t_2 is fixed. We now assume that $x(t)$ is that motion which makes L an extremum. To find the equation for $x(t)$ we introduce

$$\bar{x}(t) = x(t) + \epsilon(t) \tag{2.2.21}$$

where $\epsilon(t)$ is a small but arbitrary deviation from $x(t)$. One frequently writes

$$\bar{x}(t) - x(t) = \epsilon(t) \equiv \delta x(t) \tag{2.2.22}$$

The fact that the motion is fixed at the end points means

$$\epsilon(t_1) = \epsilon(t_2) = 0 \ . \tag{2.2.23}$$

Substituting (2.2.21) in the expressions for T and V and keeping only terms of lowest (first) order in $\epsilon(t)$ we get:

$$T = \frac{m}{2} [\dot{x} + \dot{\epsilon}]^2 = \frac{1}{2} m \dot{x}^2 + m \dot{x} \dot{\epsilon} \tag{2.2.24}$$

$$V = V(x+\epsilon) = V(x) + \epsilon\frac{\partial V}{\partial x} \,. \tag{2.2.25}$$

Thus,

$$\begin{aligned} \delta L &= \int_{t_1}^{t_2} [T(\bar{x}(t)) - V(\bar{x}(t))]\,dt - \int_{t_1}^{t_2} [T(x(t)) - V(x(t))]\,dt \\ &= \int_{t_1}^{t_2} \left[m\dot{x}\dot{\epsilon} - \epsilon\frac{\partial V}{\partial x}\right] dt \,. \end{aligned} \tag{2.2.26}$$

For an extremum we must have $\delta L = 0$. So we get:

$$\int_{t_1}^{t_2} \left[m\dot{x}\frac{d\epsilon}{dt} - \epsilon\frac{\partial V}{\partial x}\right] dt = 0 \,. \tag{2.2.27}$$

Integrating the first term by parts and using (2.2.23) we find

$$\int_{t_1}^{t_2} -\epsilon(t)\left[m\ddot{x} + \frac{\partial V}{\partial x}\right] dt = 0 \,. \tag{2.2.28}$$

But $\epsilon(t)$ is arbitrary so whenever $m\ddot{x} + \partial V/\partial x$ is positive we can make $\epsilon(t)$ negative and vice versa. Thus, we can always make the integrand positive. Therefore the only way to ensure that for arbitrary $\epsilon(t)$ the integrand in (2.2.27) vanishes is to have

$$m\ddot{x} + \frac{\partial V}{\partial x} = 0 \,. \tag{2.2.29}$$

This, however, is just Newton's law of motion with which we started.

If we had carried this computation through for an arbitrary Lagrangian $\mathcal{L}$, equation (2.2.29) would read

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} - \frac{\partial \mathcal{L}}{\partial x} = 0 \,. \tag{2.2.30}$$

Either of equations (2.2.29) and (2.2.30) is known as the Euler-Lagrange equation for this system. Thus, for a general system, to find the equations of motion find the kinetic energy T and potential V and form $\mathcal{L} = T - V$. The Euler-Lagrange equation is then the equation of motion.

To obtain Hamilton's equations (2.2.11), (2.2.12) we transform variables from x , $\dot{x}$ to x , p. Equation (2.2.2) defining the momentum p may be written more generally as

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}} \,. \tag{2.2.31}$$

We then define the Hamiltonian function $H(x,p)$ by the Legendre transformation

$$H(x,p) = \frac{\partial \mathcal{L}}{\partial \dot{x}}\dot{x} - \mathcal{L} \,, \tag{2.2.32}$$

where $\dot{x}$ is replaced by p as obtained from (2.2.31). In our example (2.2.31) yields:

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} \tag{2.2.33}$$

or

$$\dot{x} = \frac{p}{m} \,. \tag{2.2.34}$$

Then,

$$\begin{aligned} H(x,p) &= \frac{\partial \mathcal{L}}{\partial \dot{x}}\dot{x} - \mathcal{L} \\ &= m\dot{x}\dot{x} - \left[\frac{1}{2}m\dot{x}^2 - V(x)\right] \,, \end{aligned} \tag{2.2.35}$$

or finally

$$H(x,p) = \frac{p^2}{2m} + V(x) \,. \tag{2.2.36}$$

This result agrees, of course, with our previous equation (2.2.8).

In the next section we generalize these results to a system of many particles.

2.3 Lagrangian and Hamiltonian Formulation

In classical mechanics, as we stated before, the state of a system at a given time t is specified by giving the values of a complete set of dynamical variables at time t . A complete set is one from which all other dynamical variables at time t may be calculated. For example we may specify all generalized coordinate $q_1, \ldots, q_N$ and the corresponding velocities $\dot{q}_1, \cdots, \dot{q}_N$. In the Lagrange formulation there is then one scalar function $\mathcal{L}(q_1, \ldots, q_N, \dot{q}_1, \ldots, \dot{q}_N)$ from which the equations of motion for all coordinates are determined by the so-called Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_r}\right) - \frac{\partial \mathcal{L}}{\partial q_r} = 0 \,. \tag{2.3.37}$$

These equations result from finding an extremum of the action integral Thus, (2.3.37) follows from

$$\delta \int_{t_1}^{t_2} \mathcal{L} \, dt = 0 \,. \tag{2.3.38}$$

Where the variations at the end points vanish. By introducing a Legendre transformation one arrives at Hamilton's principle. Therefore, we define the generalized momenta p_r conjugate to q_r according to

$$p_r \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}_r} \tag{2.3.39}$$

and the Hamiltonian function H by the Legendre transformation

$$H = H(q_1, \ldots, q_N, p_1, \ldots, p_N) = \sum_{r=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{q}_r} \dot{q}_r - \mathcal{L} \, . \tag{2.3.40}$$

The $\dot{q}_r$ have to be eliminated in the Hamiltonian H and replaced by the p_r using (2.3.39). The resulting equations of motion are:

$$\dot{q}_r = \frac{\partial H}{\partial p_r} \quad \dot{p}_r = -\frac{\partial H}{\partial q_r} \, . \tag{2.3.41}$$

Notice that classically one may get many different Hamiltonians to describe the same system by performing so-called contact transformations about which we will say more shortly. First, some examples. Consider the Lagrangian

$$\mathcal{L}_a = \frac{1}{2} m \dot{q}^2 \, . \tag{2.3.42}$$

The equation of motion is

$$\ddot{q} = 0. \tag{2.3.43}$$

Now consider the (artificial) Lagrangian

$$\mathcal{L}_b = A e^{\beta \dot{Q}} \, . \tag{2.3.44}$$

The equation of motion is

$$\frac{d}{dt} \left(\beta A e^{\beta \dot{Q}} \right) = 0 \tag{2.3.45}$$

which yields

$$\ddot{Q} = 0 \, . \tag{2.3.46}$$

Thus, both Lagrangians give the same equation of motion. We now look at their Hamiltonians.

In the first case we get

$$p = \frac{\partial \mathcal{L}_a}{\partial \dot{q}} = m \dot{q} \tag{2.3.47}$$

and

$$H_a = \frac{1}{2} m \dot{q}^2 = \frac{p^2}{2m} \, . \tag{2.3.48}$$

This is clearly the total energy. In the second case we get

$$P = \frac{\partial \mathcal{L}_b}{\partial \dot{Q}} = \beta A e^{\beta \dot{Q}} \tag{2.3.49}$$

so that

$$\dot{Q} = \frac{1}{\beta} \ln \frac{P}{\beta A} \tag{2.3.50}$$

and

$$H_b = A\left[\beta\dot{Q} - 1\right] e^{\beta\dot{Q}} \tag{2.3.51}$$

or, after substituting for $\dot{Q}$

$$H_b = \left[\ln\frac{P}{\beta A} - 1\right]\frac{P}{\beta} \ . \tag{2.3.52}$$

Although this has the dimensions of an energy, it does not represent the total energy of a free particle. The Hamilton equations of motion are:

(a)

$$\begin{aligned} \dot{p} &= -\frac{\partial H_a}{\partial q} = 0 \\ \dot{q} &= \frac{\partial H_a}{\partial p} = \frac{p}{m} \end{aligned}$$

(b)

$$\begin{aligned} \dot{P} &= -\frac{\partial H_b}{\partial Q} = 0 \\ \dot{Q} &= \frac{\partial H_b}{\partial P} = \frac{1}{\beta}\ln\frac{P}{\beta A} \ . \end{aligned}$$

For classical mechanics it clearly does not matter much which Hamiltonian we use, but for quantum mechanics it does. Therefore, we agree henceforth to choose the Hamiltonian H such that H coincides with the total energy of the system.

2.4 Contact Transformations: Hamilton-Jacobi Theory

In classical mechanics, given the set of variables $\{q_i, p_i\}$ we can introduce new variables

$$\begin{aligned} Q_i &= Q_i(q_r, p_r) \\ P_i &= P_i(q_r, p_r) \ . \end{aligned} \tag{2.4.53}$$

Of particular interest are those transformations that preserve the form of Hamilton's equations. Thus, there must exist a function $K(Q_i, P_i)$ such that the equations of motion read

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} \ , \ \dot{P}_i = -\frac{\partial K}{\partial Q_i} \ . \tag{2.4.54}$$

In this case, such transformations (2.4.54) are called contact or canonical. Since Q_i , P_i satisfy (2.4.54) they must also satisfy a variational principle (2.3.38). But,

$$\mathcal{L} = \sum_i p_i \dot{q}_i - H \tag{2.4.55}$$

and

$$\delta \int_{t_1}^{t_2} \left[\sum_i p_i \dot{q}_i - H(p_i, q_i) \right] dt = 0 \tag{2.4.56}$$

as well as

$$\delta \int_{t_1}^{t_2} \left[\sum_i P_i \dot{Q}_i - K(P_i, Q_i) \right] dt = 0 \ . \tag{2.4.57}$$

Thus, the two integrands can differ at most by a total differential dF/dt since then

$$\int_{t_1}^{t_2} \frac{dF}{dt} dt = F(2) - F(1) \tag{2.4.58}$$

and the variation

$$\delta \int_{t_1}^{t_2} \frac{dF}{dt} dt = F(2) - F(1) = 0 \ . \tag{2.4.59}$$

The function F is called the generating function of the transformation. Naively, one expects F to be a function of the $4N + 1$ variables q_i , p_i , Q_j , P_j , t. However, due to the connecting equations (2.4.54) only $2N + 1$ are independent. For example

$$F = F_1(q_i, Q_j, t) \ . \tag{2.4.60}$$

Since the integrands differ only by dF/dt we have

$$\sum_i p_i \dot{q}_i - H = \sum_i P_i \dot{Q}_i - K + \frac{dF_1}{dt} \ . \tag{2.4.61}$$

But

$$\frac{dF_1}{dt} = \sum_i \frac{\partial F_1}{\partial q_i} \dot{q} + \sum_i \frac{\partial F_1}{\partial Q_i} \dot{Q}_i + \frac{\partial F_1}{\partial t} \ . \tag{2.4.62}$$

Since the q_i and Q_i are independent variables we get from (2.4.61) and (2.4.62) that

$$p_i = \frac{\partial F_1}{\partial q_i} \tag{2.4.63}$$

$$P_i = -\frac{\partial F_1}{\partial Q_i} \tag{2.4.64}$$

$$K = H + \frac{\partial F_1}{\partial t} \; . \tag{2.4.65}$$

The equations of motion (2.4.54) in the new variables are particularly easy to solve if

$$\frac{\partial K}{\partial P_i} = \dot{Q}_i = 0 \tag{2.4.66}$$

and

$$-\frac{\partial K}{\partial Q_i} = \dot{P}_i = 0 \; . \tag{2.4.67}$$

This is most easily achieved by choosing $K = 0$. Thus, F_1 is determined by

$$\frac{\partial F_1}{\partial t} + H = 0 \quad . \tag{2.4.68}$$

In this case it is more convenient to choose a function F_2 as a function of the q_i and P_i. This generating function F_2 is related to F_1 by a Legendre transformation

$$F_2(q_i, P_i, t) = F_1(q_i, Q_i, t) + \sum P_i Q_i \quad . \tag{2.4.69}$$

It is then an easy matter to check that instead of (2.4.63) - (2.4.65) we get

$$p_i = \frac{\partial F_2}{\partial q_i} \tag{2.4.70}$$

$$Q_i = \frac{\partial F_2}{\partial P_i} \tag{2.4.71}$$

$$K = H + \frac{\partial F_2}{\partial t} \; . \tag{2.4.72}$$

Thus with $H = H(q_i, p_i)$ the equation for $K = 0$ reads:

$$H\left(q_1, \cdots, q_N; \frac{\partial F}{\partial q_1}, \cdots, \frac{\partial F}{\partial q_N}; t\right) + \frac{\partial F}{\partial t} = 0 \tag{2.4.73}$$

where we have dropped the subscript on F and used (2.4.70) to replace the p_i by $\partial F/\partial q_i$. This is the celebrated Hamilton-Jacobi equation.

Example:

$$H = \frac{p^2}{2m} \; . \tag{2.4.74}$$

Then,

$$\frac{\partial F}{\partial t} + \frac{1}{2m}\left(\frac{\partial F}{\partial q}\right)^2 = 0 \tag{2.4.75}$$

and a solution of (2.4.75) is

$$F = \alpha q - \beta t \tag{2.4.76}$$

with

$$\beta = \frac{\alpha^2}{2m} . \tag{2.4.77}$$

Before proceeding we do some more formal manipulations that will give us a basis later for a heuristic "derivation" of Schrödinger's equation.

Consider the Hamilton-Jacobi equation (2.4.73) and call S a solution of it. Equation (2.4.73) is a first order partial differential equation in $N+1$ variables. However if S is a solution so is $S+\alpha$ for any constant α. Thus S contains only N constants as far as transformations are concerned. Furthermore,

$$\dot{P}_i = 0 . \tag{2.4.78}$$

Thus the P_i are constants. Hence,

$$S = S(q_1 \cdots q_N, \alpha_1 \cdots \alpha_N, t) \tag{2.4.79}$$

is a solution where

$$\begin{aligned} P_i &= \alpha_i \\ p_i &= \frac{\partial S(q_i, \alpha_i, t)}{\partial q_i} \end{aligned} \tag{2.4.80}$$

and

$$Q_i = \beta_i = \frac{\partial S(q_i, \alpha_i, t)}{\partial \alpha_i} . \tag{2.4.81}$$

We can now invert these equations (2.4.79) and get:

$$q_i = q_i(\alpha_r, \beta_r, t) . \tag{2.4.82}$$

In our example

$$S = -\frac{\alpha^2 t}{2m} + \alpha q \tag{2.4.83}$$

and

$$Q = \beta = \frac{\partial S}{\partial \alpha} = -\frac{\alpha t}{m} . \tag{2.4.84}$$

So

$$q = \beta + \frac{\alpha t}{m} . \tag{2.4.85}$$

Thus, solving the Hamilton-Jacobi equation

$$H\left(q_1, \cdots, q_N; \frac{\partial S}{\partial q_1}, \cdots, \frac{\partial S}{\partial q_N}; t\right) + \frac{\partial S}{\partial t} = 0 \tag{2.4.86}$$

gives a solution of the dynamical problem. Note that we can always write

$$S(q_i, \alpha_i, t) = W(q_i, \alpha_i) - \alpha_1 t \,. \tag{2.4.87}$$

Then (2.4.86) becomes

$$H\left(q_1, \cdots, q_N; \frac{\partial W}{\partial q_1}, \cdots, \frac{\partial W}{\partial q_N}; t\right) = \alpha_1 \,. \tag{2.4.88}$$

This equation is t-independent. We can separate the equation even further by writing

$$W = \sum_i W_i(q_i, \alpha_1 \cdots \alpha_N) \,. \tag{2.4.89}$$

Then we get N equations

$$H\left(q_i, \frac{\partial W}{\partial q_i}, \alpha_1 \cdots \alpha_N\right) = \alpha_1 \,. \tag{2.4.90}$$

These are first order ordinary differential equations. The momenta p_i are still given by the equations of the canonical transformations

$$p_i = \frac{\partial W(q_i, \alpha_1 \cdots \alpha_N)}{\partial q_i} \,. \tag{2.4.91}$$

If the motion is periodic, then the action integral

$$J_i = \oint p_i dq_i = \oint \frac{\partial W_i}{\partial q_i} dq_i \tag{2.4.92}$$

is a function only of $\alpha_1 \cdots \alpha_N$. Thus, J_i is a constant. The rules for quantization given by Bohr were extended by Sommerfeld to cover such periodic motions according to

$$\oint p_i dq_i = nh = 2\pi\hbar n \,. \tag{2.4.93}$$

Here we have introduced the new constant $\hbar = h/2\pi$. This constant will occur more often than the simple Planck's constant. The rules stated in equation (2.4.93) are the so-called Bohr-Sommerfeld quantization rules. In order to understand where they come from we now discuss the meaning of the J_i . As an example we use our old friend the simple harmonic oscillator.

$$H = \frac{p^2}{2m} + \frac{k}{2}q^2 \,. \tag{2.4.94}$$

Setting $p = \frac{\partial S}{\partial q}$ we get the equation

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{k}{2}q^2 + \frac{\partial S}{\partial t} = 0 \,. \tag{2.4.95}$$

Now we set

$$S(q,\alpha,t) = W(q,\alpha) - \alpha t \ . \tag{2.4.96}$$

Then,

$$\frac{1}{2m}\left(\frac{\partial W}{\partial q}\right)^2 + \frac{k}{2}q^2 = \alpha \tag{2.4.97}$$

so that

$$W = \sqrt{mk}\int\sqrt{\frac{2\alpha}{k} - q^2 dq} \tag{2.4.98}$$

and

$$S = \sqrt{mk}\int\sqrt{\frac{2\alpha}{k} - q^2 dq} - \alpha t \ . \tag{2.4.99}$$

Also q is given by

$$\beta = \frac{\partial S}{\partial \alpha} = \sqrt{\frac{m}{k}}\int\frac{dq}{\sqrt{\frac{2\alpha}{k} - q^2}} - t \tag{2.4.100}$$

or

$$\beta = -t - \sqrt{\frac{m}{k}}\cos^{-1}[\sqrt{\frac{k}{2\alpha}}\,q] \ . \tag{2.4.101}$$

Thus

$$q = \sqrt{\frac{2\alpha}{k}}\cos\omega(t+\beta) \tag{2.4.102}$$

where

$$\omega = \sqrt{\frac{k}{m}} \ . \tag{2.4.103}$$

To get α and β we have to impose initial conditions. For example if at $t = 0$, $p = 0$, and $q = q_0$ then $\beta = 0$ and $\alpha = k/2q_0^2$. Thus, α = total initial energy. Then, finally

$$q = q_0 \cos\omega t \ , \tag{2.4.104}$$

a result we could have obtained by much more elementary means.

2.5 Interpretation of Action-Angle Variables

In this case

$$J = \oint p\,dq = \oint \frac{\partial W(q,\alpha)}{\partial q}\,dq$$
$$= \sqrt{mk}\oint \sqrt{\frac{2\alpha}{k} - q^2}\,dq\ . \qquad (2.5.105)$$

Let

$$q = \sqrt{\frac{2\alpha}{k}}\sin\theta\ . \qquad (2.5.106)$$

Then,

$$J = 2\alpha\sqrt{\frac{m}{k}}\int_0^{2\pi}\cos^2\theta d\theta = 2\pi\alpha\sqrt{\frac{m}{k}} = \frac{2\pi\alpha}{\omega}\ . \qquad (2.5.107)$$

Therefore,

$$\alpha = H = \frac{J\omega}{2\pi} = J\nu \qquad (2.5.108)$$

so that

$$\frac{\partial H}{\partial J} = \nu\ . \qquad (2.5.109)$$

This last result is of more general validity than the derivation indicates.

Now from the correspondence principle we have for $\nu_{classical}$ a function of E as the quantum number $m \to \infty$

$$\frac{E_{m+n} - E_n}{n} \to h\nu_{classical}\ . \qquad (2.5.110)$$

Thus,

$$\frac{\partial E}{\partial n} \to h\nu_{classical} = 2\pi\hbar\nu_{classical} \qquad (2.5.111)$$

and we get as a quantization rule

$$\int_{E_{min}}^{E_{max}} \frac{dE}{\nu_{classical}(E)} = 2\pi n\hbar\ . \qquad (2.5.112)$$

But,

$$\frac{\partial H}{\partial J_i} = \frac{\partial E}{\partial J_i} = \nu_i \qquad (2.5.113)$$

so

$$\frac{dE}{\nu_i} = dJ_i\ . \qquad (2.5.114)$$

Thus,

$$\oint dJ_i = \oint p_i dq_i = 2\pi n\hbar\ . \qquad (2.5.115)$$

This is one way to arrive at Bohr-Sommerfeld quantization.

2.6 Hydrogen Atom: Bohr-Sommerfeld Quantization

As an example of the use of this technique we work out the hydrogen atom including elliptical orbits. The Hamiltonian is:

$$H = \frac{p^2}{2m} - \frac{e^2}{r} = \frac{1}{2m}\left[p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2\theta}\right] - \frac{e^2}{r} \tag{2.6.116}$$

where now e is measured in e.s.u. The quantization is given by

$$\oint p_i dq_i = nh = 2\pi n\hbar \ . \tag{2.6.117}$$

Also,

$$\begin{aligned} p_r &= m\dot{r} \\ p_\theta &= mr^2\dot{\theta} \\ p_\varphi &= mr^2 \sin^2\theta \, \dot{\varphi} \ . \end{aligned} \tag{2.6.118}$$

The Hamilton-Jacobi equation for W is

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W}{\partial \theta}\right)^2 + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial W}{\partial \varphi}\right)^2\right] - \frac{e^2}{r} = \alpha_1 = E. \tag{2.6.119}$$

Separating the variables

$$W = W_r(r) + W_\theta(\theta) + W_\varphi(\varphi) \tag{2.6.120}$$

we get

$$\begin{aligned} \frac{\partial W_\varphi}{\partial \varphi} &= \alpha_\varphi = \text{constant} \\ \left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + \frac{\alpha_\varphi^2}{\sin^2\theta} &= \alpha_\theta^2 \\ \left(\frac{\partial W_r}{\partial r}\right)^2 + \frac{\alpha_\theta^2}{r^2} &= 2m\left(E + \frac{e^2}{r}\right) \ . \end{aligned} \tag{2.6.121}$$

These last three equations are three conservation statements. The first gives:

$$p_\varphi = \alpha_\varphi \quad \text{conservation of } p_\varphi \ . \tag{2.6.122}$$

The second gives:

$$p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta} = \alpha_\theta^2 \ . \tag{2.6.123}$$

This is just conservation of total angular momentum and follows from the fact that

$$H = \frac{1}{2m}\left[p_r^2 + \frac{L^2}{r^2}\right] - \frac{e^2}{r} \tag{2.6.124}$$

so that

$$L^2 = p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta} \tag{2.6.125}$$

as stated.

We now apply the quantization.

$$J_\varphi = \oint p_\varphi d\varphi = \oint \alpha_\varphi d\varphi = 2\pi\alpha_\varphi = 2\pi\, m'\hbar \tag{2.6.126}$$

so that the constant α_φ is quantized and given by

$$\alpha_\varphi = m'\hbar\,. \tag{2.6.127}$$

Then,

$$J_\theta = \oint \sqrt{\alpha_\theta^2 - \frac{\alpha_\theta^2}{\sin^2\theta}}d\theta\,. \tag{2.6.128}$$

To evaluate this integral we use a trick.

Since the equations for generalized coordinates do not involve the time t explicitly we have for the kinetic energy T

$$2T = \sum p_i \dot{q}_i \tag{2.6.129}$$

so that in polar coordinates

$$2T = p_r\dot{r} + L\dot{\psi} = p_r\dot{r} + p_\theta\dot{\theta} + p_\varphi\dot{\varphi}\,. \tag{2.6.130}$$

Here ψ = angle of azimuth of the particle in its orbit. (See figure 2.1.) Thus,

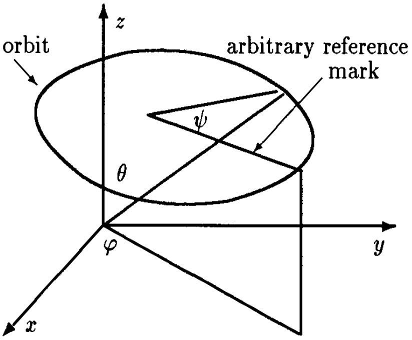

Figure 2.1: Orbit of electron in a hydrogen atom.

$$p_\theta\, d\theta = L\, d\psi - p_\varphi\, d\varphi\,. \tag{2.6.131}$$

From (2.6.131) we get:

$$J_\theta = \oint L\, d\psi - \oint p_\varphi\, d\varphi\,. \tag{2.6.132}$$

Now as θ goes through a complete cycle ψ and φ go through 2π. Also, as we already saw,

$$L = \alpha_\theta = \text{constant} \tag{2.6.133}$$

and

$$p_\varphi = \alpha_\varphi\,. \tag{2.6.134}$$

Therefore,

$$J_\theta = 2\pi(\alpha_\theta - \alpha_\varphi) = 2\pi(l - m')\hbar\,. \tag{2.6.135}$$

Hence,

$$J_r = \oint \left[2mE + \frac{2me^2}{r} + \frac{(J_\theta + J_\varphi)^2}{4\pi^2 r^2}\right]^{1/2} dr = 2\pi(n-l)\hbar\,. \tag{2.6.136}$$

We evaluate this integral by contour integration. Now,

$$J_\theta + J_\varphi = 2\pi l\hbar\,. \tag{2.6.137}$$

Then,

$$J_r = \oint \left[2mE + \frac{2me^2}{r} - \frac{l^2\hbar^2}{r^2}\right]^{1/2} dr\,. \tag{2.6.138}$$

The motion is bounded only if $E < 0$. The maximum and minimum values of r are then obtained from

$$-2m|E| + \frac{2me^2}{r} - \frac{l^2\hbar^2}{r^2} = 0\,. \tag{2.6.139}$$

This gives:

$$r_{1,2} \equiv r_{\text{max,min}} = \frac{2me^2 \pm \sqrt{4m^2e^4 - 8m|E|l^2\hbar^2}}{4m|E|}\,. \tag{2.6.140}$$

Considered as a function of z, the function

$$\left[-2m|E| + \frac{2me^2}{z} - \frac{l^2\hbar^2}{z^2}\right]^{1/2}$$

has branch points at r_1 and r_2 and poles at 0 and ∞. (See figure 2.2.)

The integral for J_r is given by the contour shown. However, by considering this contour as enclosing its external rather than its internal part we get the same result if we integrate in the reverse direction (clockwise). Then,

$$\begin{aligned} J_r &= 2\pi i \times (\text{residue at 0 and} \infty) \\ &= 2\pi(me^2/\sqrt{2m|E|} - l\hbar) \\ &= 2\pi(n-l)\hbar\,. \end{aligned} \tag{2.6.141}$$

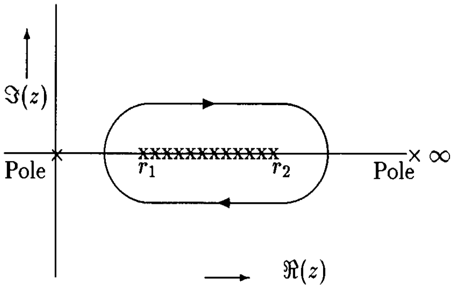

Figure 2.2: Contour used to evaluate J_r.

Thus,

$$|E| = -E = \frac{1}{2}\frac{me^4}{n^2\hbar^2} \tag{2.6.142}$$

or

$$E = -\frac{1}{2}mc^2\frac{e^4}{\hbar^2c^2}\frac{1}{n^2} = -\frac{1}{2}mc^2\alpha^2\frac{1}{n^2}\,. \tag{2.6.143}$$

So, the Rydberg constant is given by

$$R_y = \frac{1}{2}mc^2\alpha^2/\hbar \tag{2.6.144}$$

just as we found before.

On the other hand, the procedure we just used is very clumsy and applies only to periodic motions. We now look for some way to generalize this procedure. To do this we study the connection between geometrical and physical optics to obtain a generalization of classical mechanics.

2.7 The Schrödinger Equation

Consider a one-dimensional system with Hamiltonian

$$H = T + V = \frac{p^2}{2m} + V \tag{2.7.145}$$

which is:
1) The total energy;
2) A constant of the motion.
As previously stated we only consider such Hamiltonians.

Now the principal function S and the characteristic function W are related by

$$S(q,p,t) = W(q,p) - Et\,. \tag{2.7.146}$$

Furthermore,

$$p = \frac{\partial S}{\partial q} \ . \tag{2.7.147}$$

This means that

$$p = \frac{\partial S}{\partial q} = \frac{\partial W}{\partial q} \ . \tag{2.7.148}$$

Therefore, since the Hamiltonian H equals the total energy E the Hamilton-Jacobi equation for Hamilton's characteristic function reads:

$$\frac{1}{2m}\left(\frac{\partial W}{\partial q}\right)^2 + V(q) = E \ . \tag{2.7.149}$$

This equation tells us how to determine the actual motion of a particle if we are given W. From (2.7.149) we see that

at $t = 0$ $S = W$
at $t = 1$ the surface $S = 0$ coincides with $W = E$
at $t = 2$ the surface $S = E$ coincides with $W = 2E$

and so on.

Thus, as time progresses, a fixed surface $S =$ constant moves over the surfaces $W =$ constant. Therefore, we can say that S describes a wave motion in coordinate space. What is the speed of propagation of these waves? In going from (x, t) to $(x + dx, t + dt)$ the change in S is

$$\delta S = \frac{\partial S}{\partial x} \cdot \delta x + \frac{\partial S}{\partial t} \cdot \delta t \ . \tag{2.7.150}$$

If this is just due to propagation with a velocity u during the time dt then we have

$$\delta S = 0 \tag{2.7.151}$$

and

$$\delta x = u \delta t \ . \tag{2.7.152}$$

Thus,

$$\frac{\partial S}{\partial t} + u\frac{\partial S}{\partial x} = 0 \tag{2.7.153}$$

or

$$\frac{\partial S}{\partial t} + u \cdot p = 0 \ . \tag{2.7.154}$$

Hence,

$$-\frac{\partial S}{\partial t} = u \cdot p = mvu \tag{2.7.155}$$

where v is the particle velocity. Now,

$$\frac{\partial S}{\partial t} = -E \tag{2.7.156}$$

and

$$E = \frac{1}{2}mv^2 + V \tag{2.7.157}$$

so that

$$v = \sqrt{\frac{2(E-V)}{m}} \tag{2.7.158}$$

and we get

$$u = \frac{E}{\sqrt{2m(E-V)}} \ . \tag{2.7.159}$$

Also the Hamilton-Jacobi equation can be rewritten in terms of u to read

$$\left(\frac{\partial S}{\partial x}\right)^2 = \frac{1}{u^2}\left(\frac{\partial S}{\partial t}\right)^2 \ . \tag{2.7.160}$$

This already has some of the appearance of a wave equation.

It is therefore very plausible for us to think that classical mechanics is some sort of approximation to a wave theory. Hence we write down the simplest possible wave equation

$$\frac{\partial^2 \Psi}{\partial x^2} = \frac{1}{u^2}\frac{\partial^2 \Psi}{\partial t^2} \tag{2.7.161}$$

where Ψ is to describe this quantum mechanical wave whose interpretation we must look for later. Now since the velocity u depends only on x we can separate out the time by writing

$$\Psi(x,t) = \psi(x)e^{-iEt/\hbar} \ . \tag{2.7.162}$$

The minus sign in the exponential is conventional. The choice of $E/\hbar$ as the coefficient of t is motivated by Planck's law since $E/\hbar$ is an angular frequency. Substituting (2.7.162) into (2.7.161) and using (2.7.159) we find

$$\frac{\partial^2 \psi}{\partial x^2} = -\frac{2m(E-V)}{\hbar^2}\psi \tag{2.7.163}$$

or

$$-\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial x^2} + V\psi = E\psi \ . \tag{2.7.164}$$

This is the time-independent Schrödinger equation.

To recover the time dependence multiply $\psi(x)$ by $\exp(-iEt/\hbar)$. Thus, the time-dependent $\Psi(x,t)$ satisfies:

$$-\frac{\hbar^2}{2m}\frac{\partial^2 \Psi}{\partial x^2} + V\Psi = i\hbar\frac{\partial \Psi}{\partial t} \ . \tag{2.7.165}$$

This is the time-dependent Schrödinger equation.

The steps above do not in any way constitute a derivation of these two equations, they are simply plausibility arguments. These equations describe a new set of physical laws and their validity derives from experimental tests of the effects they predict. Our task for most of the rest of this book will be to study the physical interpretation and meaning of these equations.

2.8 Problems

2.1 Find the Lagrangian for an harmonic oscillator. Use the definition of conjugate momentum to find p and H.

2.2 Repeat problem 2.1 for the simple pendulum. Interpret the momentum p conjugate to the angle variable θ.

2.3 Use Bohr-Sommerfeld quantization to calculate the energy levels of a one-dimensional simple harmonic oscillator.

2.4 Use Bohr-Sommerfeld quantization to calculate the energy levels of a particle confined to a box of length L. For simplicity assume this is a "one-dimensional box".

2.5 Suppose a gyroscope has a magnetic moment $\vec{\mu}$ proportional to its angular momentum $\mathbf{L}$ according to

$$\vec{\mu} = M\mathbf{L}\,.$$

The potential energy due to placing the gyroscope in a magnetic field $\mathbf{B}$ is $V = -\vec{\mu}\cdot\mathbf{B}$. Assume $\mathbf{B}$ is constant and derive the equation of motion for $\mathbf{L}$. Show that the gyroscope precesses with the angular Larmor frequency

$$\omega_L = MB\,.$$

2.6 The system of quantization proposed by Bohr in 1913 is not applicable to all systems. To what general kinds of physical systems is Bohr's procedure applicable? For what kinds of systems is it not applicable.

2.7 Consider the time-dependent Schrödinger equation (2.7.165) and put

$$\Psi = Ae^{iS/\hbar}$$

where $A =$ constant. Show that in the limit as $\hbar \to 0$ equation (2.7.164) reduces to the Hamilton-Jacobi equation (2.7.149).

2.8 In problem 2.7 set $S = W - Et$ and let $W = W_0 + \hbar W_1 + \hbar^2 W_2 + \cdots$ for the case of a one-dimensional Schrdinger equation. Find the equations for W_0 and W_1 and solve the equation for W_0. This is the so-called *Wentzel-Kramers-Brillouin* or WKB approximation.

Bibliography

[2.1] W.R. Hamilton - *Collected Papers Vol. II*, Cambridge University Press (1980), pp. 103-211.

[2.2] Two of the standard reference on classical mechanics are:
H. Goldstein - *Classical Mechanics*, Addison-Wesley Publishing Co., Inc., Reading, Mass., U.S.A. (1959).
L.D. Landau and E.M. Lifschitz - *Mechanics*, Pergamon Press - U.S. distributors, Addison-Wesley Publishing Co., Inc., Reading, Mass. (1960).

Chapter 3

Elementary Systems

3.1 Introduction

In the last chapter (section 2.7) we gave a heuristic derivation of the Schrödinger equation. It is important to note that no valid derivation of this equation from classical mechanics is possible since it represents a new law of physics extending beyond classical mechanics. From this point of view it is clear that section 2.7 was only in the nature of a plausibility argument and nothing was missed if you skipped it.

Just as in classical mechanics one does not derive Newton's laws but simply starts with them, so here we shall simply start with the Schrödinger equation. The purpose of this chapter is to start familiarizing ourselves with this equation. We begin by considering the simplest case of a free particle and consider the solutions of the corresponding Schrödinger equation. We next obtain a conservation law on the basis of the Schrödinger equation. This law provides a guide for obtaining Born's probability interpretation for the solutions to the Schrödinger equation. This interpretation is then examined by considering Young's double slit experiment. We then proceed to look at some properties of a wave packet solution to see how it describes a particle. There we also introduce the concept of group and phase velocity.

To see how to make the transition from classical to quantum mechanics we go on to consider some purely formal analysis as well as mathematical relations. These lead us to some formal rules for quantizing. Sometimes, however, some of these rules can lead to ambiguities and we discuss some of the most prominent ambiguities. Finally we consider the Hamiltonian function for the very important electromagnetic interaction. The results of this last section will not be used until Chapter 15 and may thus be skipped on a first reading.

3.2 Plane Wave Solutions

As a first look at the Schrödinger equation consider the case for a particle in free space, where no forces are acting on it or the potential $V = 0$. Then

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi\,. \tag{3.2.1}$$

A solution is

$$\Psi_{\mathbf{p}}(\mathbf{r},t) = A\exp\left[\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{r} - Et)\right] \tag{3.2.2}$$

where

$$E = \frac{\mathbf{p}^2}{2m}\ . \tag{3.2.3}$$

This solution describes a wave of frequency

$$\omega = \frac{E}{\hbar}\ \text{(Planck's relation)} \tag{3.2.4}$$

and a wavelength

$$\lambda = \frac{2\pi\hbar}{p}\ \text{(de Broglie's relation)}. \tag{3.2.5}$$

Thus, two of our previous quantum mechanical results are automatically given by the Schrödinger equation.

The solution (3.2.2) is, however, physically not very realistic since according to it a physical particle has associated with it a wave uniformly spread out through all space. Fortunately this is not a defect of the theory but rather of our treatment. Since equation (3.2.1) is linear, the most general solution of it is an arbitrary linear superposition of solutions of the form (3.2.2). Thus, we can get solutions of the form

$$\Psi(\mathbf{r},t) = \int A(\mathbf{p})\exp\left[\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{r} - \frac{\mathbf{p}^2}{2m}t)\right]\,d^3p\,. \tag{3.2.6}$$

As we shall see this can describe a localized wave and is called a *wave packet.* We shall now try to give a physical interpretation of the wavefunction.

3.3 Conservation Law for Particles

So far, we have stressed the similarity between light and particles. But even quantum mechanically there are some very important differences. Light may be absorbed or emitted. Thus, the number of photons changes with time and no conservation law for the number of photons holds. On the other hand, as long as we are in a non-relativistic regime, particles cannot be destroyed or created. Of course particles can form bound states as when an electron and a proton combine to form an hydrogen atom. However, we then still have two particles. Thus, in general, except for such relativistic effects as pair creation or annihilation, the number of particles is conserved. So, in a non-relativistic theory of quantum mechanics we should have conservation of particles. If we consider the Schrödinger equation, (2.7.165), it is easy to derive a conserved density and current. We have

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi \,. \tag{3.3.7}$$

If we complex conjugate this expression we get

$$-i\hbar\frac{\partial\Psi^*}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi^* + V\Psi^* \,. \tag{3.3.8}$$

Multiplying (3.3.7) by $(-i/\hbar)\Psi^*$ and (3.3.8) by $(i/\hbar)\Psi$ and adding we get

$$\frac{\partial\Psi^*}{\partial t}\Psi + \Psi^*\frac{\partial\Psi}{\partial t} = \frac{i\hbar}{2m}\left(\Psi^*\nabla^2\Psi - \Psi\nabla^2\Psi^*\right) \,. \tag{3.3.9}$$

This is of the form

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0 \tag{3.3.10}$$

where

$$\rho = \Psi^*\Psi \tag{3.3.11}$$

and

$$\mathbf{j} = -\frac{i\hbar}{2m}(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*) \,. \tag{3.3.12}$$

Thus, we have an equation of continuity. Integrating over a fixed volume V and calling

$$R = \int_V \rho\, d^3x \tag{3.3.13}$$

we obtain, by using the equation of continuity and the divergence theorem, that

$$\frac{dR}{dt} = -\frac{i\hbar}{2m}\int_S(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*)\, d\mathbf{A} = \int_S \mathbf{j}\cdot d\mathbf{A} \tag{3.3.14}$$

where the surface S surrounds the volume V. This suggests that the rate of change of R in the volume V in (3.3.13) is given by the flux of the current $\vec{j}$ through the surface S surrounding V. It is therefore reasonable to associate the conserved quantity R with the number of particles. For only one particle $R = 1$ if V is all of space. Hence ρ is a density for the particle. However, classically a particle is at some point or it is not; it is not all smeared out, so ρ cannot be a matter density. The most reasonable interpretation is that ρ is a probability density . This gives us an interpretation of ρ and hence of Ψ.

So we interpret Ψ to be a probability amplitude . One further comment: In the derivation of (3.3.9) we assumed that the potential V is real. This is true, of course, for classical potentials and remains true in quantum mechanics. It is also crucial in our derivation and will be examined later (see problem 3.2).

To justify our interpretation further we consider some experimental results. Very early in the history of quantum mechanics experiments were performed to test experimentally the wave nature of matter. The classic experiments which we have already discussed, are those of Davisson and Germer. Thomson and Rupp [3.1], independently used powdered crystals and obtained the analogue of Debye-Scherrer patterns. For a modern version see C. Jönsen [3.2]. We now consider a slightly idealized experiment which was not performed until much later.

3.4 Young's Double Slit Experiment

The experimental set up is well known and is shown in figure 3.1. The intensity distribution of a wave is given by $|\Psi|^2$, the square of its amplitude. This is also found to be the case for a high intensity beam of particles. On the other hand, when we decrease the beam intensity to the point where we can see individual particles arriving at the screen, we find that they arrive at specific points and not spread out all over the screen. The point at which any given particle arrives is completely unpredictable. However, $|\Psi|^2$ still gives information, in fact, the probability density. Thus, we find experimentally that $|\Psi|^2\,dx$ is the probability that the particle arrives at a point lying between x and $x+dx$. Thus, our interpretation of Ψ as a probability amplitude is again justified.

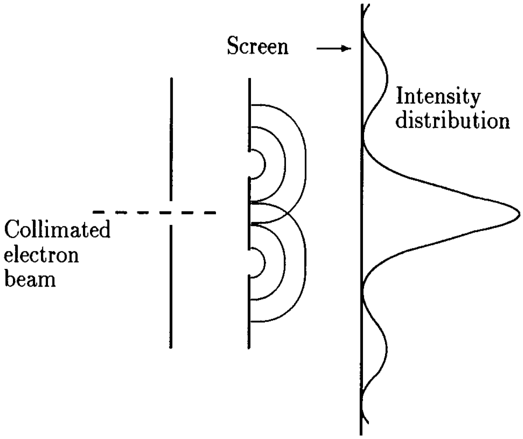

Figure 3.1: Young's double slit experiment.

Incidentally it is also possible to do this experiment with light by decreasing the intensity to the point where individual photons are observed.

3.5 The Superposition Principle and Group Velocity

A fundamental property of the Schrödinger equation is that it is linear in Ψ. This has the following important consequence. If Ψ_1 and Ψ_2 are two different solutions, then $\lambda_1\Psi_1 + \lambda_2\Psi_2$ is also a solution if λ_1 and λ_2 are constants. This is known as the superposition principle.

As an example of the use of this principle we now construct and examine a wave packet. We already wrote down the equation for a wave packet equation (3.2.6). For simplicity we work in one space dimension.

A plane monochromatic wave is given by

$$u_k(t,x) = e^{i(kx-\omega t)} \tag{3.5.15}$$

where

$$\hbar\omega = \frac{(\hbar k)^2}{2m} \tag{3.5.16}$$

for a free particle. Also, the wavelength is given by

$$\lambda = \frac{2\pi}{k}\,. \tag{3.5.17}$$

In general $\omega = \omega(k)$. Forming a superposition of such plane waves we build up a wave packet

$$\Psi(x,t) = \int_{-\infty}^{\infty} f(k) e^{i[kx-\omega(k)t]}\, dk\,. \tag{3.5.18}$$

We want to find out where this wave is concentrated in x-space and with what velocity the peak of the wave travels. This is known as the group velocity.

To do this we assume $f(k)$ is a smooth well-behaved function concentrated in a region Δk about $k = k_0$. We further assume that in this region $\omega(k)$ may be expanded in a power series about k_0.

$$\omega(k) = \omega_0 + (k-k_0)\left.\frac{d\omega}{dk}\right|_{k_0} + \cdots \tag{3.5.19}$$

where we have written ω_0 for $\omega(k_0)$. Substituting this expression into (3.5.18) we get

$$\Psi(x,t) \approx e^{i[k_0x-\omega_0t]}\int_{k_0-\Delta k}^{k_0+\Delta k} f(k)\exp\left[i(k-k_0)(x-\left.\frac{d\omega}{dk}\right|_0 t)\right] dk\,. \tag{3.5.20}$$

This expression is of the form

$$\Psi(x,t) \approx e^{i[k_0x-\omega_0t]}F\left(x-\left.\frac{d\omega}{dk}\right|_0 t\right)\,. \tag{3.5.21}$$

This is an envelope function F, the Fourier transform of f, multiplying a plane monochromatic wave. The phase velocity is given by the monochromatic wave and is

$$v_p = \frac{\omega_0}{k_0} \ . \tag{3.5.22}$$

The envelope function is unchanged if we replace x and t by $x + \delta x$ and $t + \delta t$ such that

$$\delta x = \left.\frac{d\omega}{dk}\right|_0 \delta t \ . \tag{3.5.23}$$

Thus, the group velocity is

$$v_g = \left.\frac{d\omega}{dk}\right|_{k=k_0} \ . \tag{3.5.24}$$

To make these considerations valid, the phase

$$\phi = kx - \omega(k)t \tag{3.5.25}$$

of the integrand in (3.5.18) must not vary too rapidly in the region about k_0, otherwise the positive and negative contributions cancel. In fact we require roughly that not more than one oscillation to occur in this region. This gives the condition

$$\Delta k \left.\frac{d\phi}{dk}\right|_0 \geq 1 \ . \tag{3.5.26}$$

But,

$$\frac{d\phi}{dk} = x - \frac{d\omega}{dk}t \approx x - \left.\frac{d\omega}{dk}\right|_0 t \ . \tag{3.5.27}$$

This shows that the centre of the packet is given by

$$x_0 = \left.\frac{d\omega}{dk}\right|_0 t \ . \tag{3.5.28}$$

Therefore,

$$\frac{d\phi}{dk} \approx x - \left.\frac{d\omega}{dk}\right|_0 t = (x - x_0) = \Delta x \ . \tag{3.5.29}$$

Thus, the packet is concentrated in a region Δx such that

$$\Delta k \Delta x \geq 1 \ . \tag{3.5.30}$$

Since $p = \hbar k$ this can be written $\Delta p \Delta x \geq \hbar$. This is an example of Heisenberg's uncertainty relation. We shall have more to say about this relation later on.

The group velocity v_g as previously calculated corresponds to the velocity of the classical particle. We can see this as follows. Classically,

$$v = \frac{dE}{dp} \ . \tag{3.5.31}$$

Now we also have

$$E = \hbar\omega \ , \quad p = \hbar k \tag{3.5.32}$$

and

$$\frac{dE}{dp} = \frac{d\omega}{dk} = v_g \tag{3.5.33}$$

as stated.

There is one more property of a plane wave that is very enlightening and fortifies our faith in the interpretation of Ψ given so far. This is the value of the current or flux for such a wave. If we take a plane wave

$$\Psi = Ae^{i(kx-\omega t)} \tag{3.5.34}$$

and use (3.3.12) to compute the current we get that

$$\begin{aligned} j &= -\frac{i\hbar}{2m}|A|^2\left[e^{-i(kx-\omega t)}\frac{d}{dx}e^{i(kx-\omega t)}e^{-i(kx-\omega t)}\frac{d}{dx}e^{-i(kx-\omega t)}\right] \\ &= \frac{\hbar k}{m}|A|^2 \ . \end{aligned} \tag{3.5.35}$$

Since the momentum $p = \hbar k$ and $p/m = v$, the velocity, we have:

$$j = v\,|A|^2 \ . \tag{3.5.36}$$

Thus, the current consists of the velocity v times the intensity $|A|^2$ of the wave.

3.6 Formal Considerations

This part consists mostly of some mathematical observations. However, these niceties will be very useful later. First we introduce the concept of an operator. As motivation let us re-examine the Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi \ . \tag{3.6.37}$$

This equation was derived for a system with a classical Hamiltonian

$$H = \frac{p^2}{2m} + V \ . \tag{3.6.38}$$

Now if we replace p by $\frac{\hbar}{i}\nabla$, then we get an operator H_{op} such that

$$H_{\text{op}}\,\Psi = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi \ . \tag{3.6.39}$$

And the time independent Schrödinger equation reads

$$H_{\text{op}}\,\psi = E\,\psi \ . \tag{3.6.40}$$

So our equation looks much more suggestive this way. In fact this procedure of replacing classical variables by operators has much more than merely formal significance.

Now suppose ψ is a function describing a given state, then $d\psi/dx$ is another function and is obtained by operating on ψ with d/dx. Clearly d/dx can operate only on functions where the first derivative exists. These functions then lie in the *domain* of d/dx. More generally if one has a procedure that assigns to every function of a set $\{\psi_1, \ldots, \psi_n\}$ a unique function of some other set $\{\phi_1, \ldots, \phi_m\}$ one says that this mapping is given by an operator and we write

$$(A\psi)_i = \phi_i \; . \tag{3.6.41}$$

The operators of interest in quantum mechanics are in a sense the simplest; they are linear. An operator A is linear if

$$A(\lambda_1\psi_1 + \lambda_2\psi_2) = \lambda_1 A\psi_1 + \lambda_2 A\psi_2 \tag{3.6.42}$$

where λ_1 and λ_2 are arbitrary constants and ψ_1 and ψ_2 are two arbitrary states.

Examples:

The operator d/dx is linear since

$$\frac{d}{dx}(\lambda_1\psi_1 + \lambda_2\psi_2) = \lambda_1\frac{d}{dx}\psi_1 + \lambda_2\frac{d}{dx}\psi_2 \; . \tag{3.6.43}$$

On the other hand the operator $\sqrt{\ }$ is not linear since

$$\sqrt{(\lambda_1\psi_1 + \lambda_2\psi_2)} \;\neq\; \lambda_1\sqrt{\psi_1} + \lambda_2\sqrt{\psi_2} \; . \tag{3.6.44}$$

Another important example of a linear operator is multiplication by a function say $f(x)$ since

$$f(x)(\lambda_1\psi_1 + \lambda_2\psi_2) = \lambda_1 f(x)\psi_1 + \lambda_2 f(x)\psi_2 \; . \tag{3.6.45}$$

We shall only consider linear operators from now on unless we expressly state otherwise.

It is also possible to have an algebra of operators. The following operations with operators are defined.

1) Scalar multiplication λA is defined by

$$(\lambda A)\psi = \lambda(A\psi) \; . \tag{3.6.46}$$

2) Addition

$$(A + B)\psi = A\psi + B\psi \; . \tag{3.6.47}$$

3) Multiplication

$$(AB)\psi = A(B\psi) \; . \tag{3.6.48}$$

These last two also satisfy distributive laws, namely

$$A+(B+C)=(A+B)+C \tag{3.6.49}$$

and

$$A\,(B\,C)=(A\,B)\,C\,. \tag{3.6.50}$$

On the other hand, the product is not commutative and in general $AB \neq BA$. For example, if $A = x$, $B = d/dx$

$$(AB)\psi = x\frac{d\psi}{dx} \tag{3.6.51}$$

whereas

$$(BA)\psi = \frac{d}{dx}\,(x\psi) = \psi + x\frac{d\psi}{dx}\,. \tag{3.6.52}$$

To express the lack of commutativity we introduce the commutator

$$[A\,,\,B] = AB - BA\,. \tag{3.6.53}$$

Thus, in the example above

$$[\frac{d}{dx}\,,\,x] = 1 \tag{3.6.54}$$

since

$$\frac{d}{dx}(x\psi) - x\frac{d\psi}{dx} = 1\cdot\psi\,. \tag{3.6.55}$$

The commutator plays a very important role in quantum mechanics. In a sense it shows to what extent the quantum operator differs from the corresponding classical variable. We will have much more to say about this later. Other familiar examples of operators are grad and curl.

3.7 Ambiguities

As we saw, the time independent Schrödinger equation could be written in the form

$$H_{op}\psi = E\psi\,. \tag{3.7.56}$$

The rule for forming H_{op} from the classical Hamiltonian H was simply to replace $\mathbf{p}$ by $\hbar/i\ \nabla$. This rule, however, is not completely unambiguous. There are at least two sources of ambiguity which we now discuss. The first of these has to do with the use of different coordinate systems and does not present any serious problems in principle. The second is due to the fact that operators do not usually commute. Here the ambiguities must be resolved partly by mathematics and partly by physics.

3.7.1 Use of Different Coordinate Systems

Consider the free particle Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} \ . \tag{3.7.57}$$

In Cartesian coordinates this reads

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) \tag{3.7.58}$$

giving a quantum mechanical operator Hamiltonian

$$H_{op} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \ . \tag{3.7.59}$$

In cylindrical coordinates (3.7.57) reads:

$$H = \frac{1}{2m}\left(p_r^2 + \frac{1}{r^2}p_\varphi^2 + p_z^2\right) \ . \tag{3.7.60}$$

If one naively replaces p_r by $\hbar/i\ \partial/\partial r$ and p_φ by $\hbar/i\ \partial/\partial\varphi$ this gives

$$H'_{op} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2}{\partial\varphi^2} + \frac{\partial^2}{\partial z^2}\right) , \tag{3.7.61}$$

whereas if one transforms (3.7.59) to cylindrical coordinates one gets:

$$H_{op} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\varphi^2} + \frac{\partial^2}{\partial z^2}\right) \ . \tag{3.7.62}$$

In fact, the latter is correct. The difference is due to the fact that we must distinguish between covariant and contravariant vectors. However, it is possible to get away without knowing this distinction by adhering to the following simple rule. *Always write all momenta in Cartesian coordinates and then make the replacements*

$$\begin{aligned} p_x &\to \frac{\hbar}{i}\frac{\partial}{\partial x} \\ p_y &\to \frac{\hbar}{i}\frac{\partial}{\partial y} \\ p_z &\to \frac{\hbar}{i}\frac{\partial}{\partial z} \ . \end{aligned} \tag{3.7.63}$$

After that, transform to the desired coordinate system. The problem will now usually become the standard one of writing the Laplacian in the appropriate coordinate system, and should not give any difficulties. It may also be useful to consult [3.4].

3.7.2 Non-Commutativity

In classical mechanics, since all dynamical variables commute, it does not matter in which order we write them as factors. This is not the case in quantum mechanics. For example,

$$\frac{1}{2m}p^2 = \frac{1}{2m}xp\,\frac{1}{x^2}\,px \tag{3.7.64}$$

classically. But this is definitely not the case in quantum mechanics. In fact, using the relation

$$[g(x)\,,p] = i\hbar\frac{dg(x)}{dx} \tag{3.7.65}$$

(see problem 3.1) we obtain

$$\begin{aligned} xp\,\frac{1}{x^2}px &= xp\,\frac{1}{x^2}xp + xp\,\frac{1}{x^2}[p,x] \\ &= xp\,\frac{1}{x}p - i\hbar xp\,\frac{1}{x^2} \\ &= x\frac{1}{x}p^2 + x[p\,\frac{1}{x}]p - i\hbar x\left(-\frac{2i\hbar}{x^3}\right) \\ &= p^2 + i\frac{\hbar}{x}p - i\frac{\hbar}{x}p - i\hbar x\left(-\frac{2i\hbar}{x^3}\right) \end{aligned} \tag{3.7.66}$$

or

$$xp\,\frac{1}{x^2}\,px = p^2 - \frac{2\hbar^2}{x^2} \neq p^2\;. \tag{3.7.67}$$

This example was, of course, artificially constructed, but unfortunately there is no rule for obtaining the correct order. This has to come from experiment, although in some cases mathematics can help because all physical observables must have corresponding hermitian operators. Sometimes this helps to eliminate some of the many possible orderings. Fortunately, in almost all cases of physical interest, the order is known. Furthermore, the classical Hamiltonian usually has an especially simple form

$$H = T + V \tag{3.7.68}$$

where

$$T = \text{kinetic energy} = \sum_i \frac{p_i^2}{2m} \tag{3.7.69}$$

and

$$V = \text{potential energy} = \frac{1}{2}\sum_{i\neq j} V_{ij}(x_i - x_j)\;. \tag{3.7.70}$$

In this case the dynamical terms can be translated directly and no difficulty occurs. There are a few very important exceptions. One is the so-called spin-orbit interaction which we will encounter much later and another is the interaction of a charged particle with an electromagnetic field. In both these cases we get terms of the form $\mathbf{p} \cdot \mathbf{A}(x)$. We discuss the electromagnetic interaction next. This discussion is classical but allows a direct translation to quantum mechanics. For further discussions see [3.5].

3.8 Interaction with an Electromagnetic Field

Classically the force in this case is the Lorentz Force

$$\mathbf{F} = q\,(\mathbf{E} + \frac{1}{c}\mathbf{v} \times \mathbf{B})\,. \tag{3.8.71}$$

The corresponding Lagrangian is

$$\mathcal{L} = T - q(\phi - \frac{1}{c}\mathbf{v} \cdot \mathbf{A}) \tag{3.8.72}$$

where

$$\mathbf{E} = -\mathrm{grad}\phi \tag{3.8.73}$$

$$\mathbf{B} = \mathrm{curl}\mathbf{A} \tag{3.8.74}$$

and

$$T = \frac{1}{2}m\mathbf{v}^2\,. \tag{3.8.75}$$

We are considering time-independent electromagnetic fields. If they are time dependent, equations (3.8.73) and (3.8.74) have to be suitably modified. The momentum $\mathbf{p}$ conjugate to $\mathbf{x}$ is given by

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{x}_i} = \frac{\partial \mathcal{L}}{\partial v_i} = mv_i + \frac{q}{c}A_i \tag{3.8.76}$$

or

$$\mathbf{p} = m\mathbf{v} + \frac{q}{c}\mathbf{A}\,. \tag{3.8.77}$$

Hence we get the Hamiltonian

$$\begin{aligned} H &= \sum_i p_i\dot{x}_i - \mathcal{L} = \mathbf{p} \cdot \mathbf{v} - \mathcal{L} \\ &= mv^2 + \frac{q}{c}\mathbf{v} \cdot A - \frac{1}{2}mv^2 + q\phi - \frac{q}{c}\mathbf{v} \cdot A \end{aligned} \tag{3.8.78}$$

or

$$H = \frac{1}{2}mv^2 + q\phi\,. \tag{3.8.79}$$

We now have to use (3.8.77) to replace $\mathbf{v}$ by $\mathbf{p}$. Then,

$$H = \frac{1}{2m}(\mathbf{p} - \frac{q}{c}\mathbf{A})^2 + q\phi\,, \tag{3.8.80}$$

or expanding

$$H = \frac{p^2}{2m} - \frac{q}{2mc}(\mathbf{p}\cdot\mathbf{A} + \mathbf{A}\cdot\mathbf{p}) + \frac{q^2}{2mc^2}\mathbf{A}^2 + q\phi\,. \tag{3.8.81}$$

In this form the Hamiltonian translates directly into H_{op} by replacing p by $\hbar/i\,\nabla$.

We are now ready to start looking at solutions of the Schrödinger equation.

3.9 Problems

3.1 a) Verify the identity

$$[AB, C] = A[B, C] + [A, C]B\,.$$

b) Using the result above and $[x, p] = i\hbar$ prove that

$$[x^2, p] = 2i\hbar x$$

and

$$[x^n, p] = ni\hbar x^{n-1}\quad.$$

Hence prove that for any function $g(x)$ analytic at the origin

$$[g(x), p] = i\hbar\frac{dg(x)}{dx}\quad.$$

3.2 Assume that the potential V is complex of the form $V = U + iW$. Show that W corresponds to a sink or source of probability.
Hint: Show that

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = \frac{2}{\hbar}W\rho\quad.$$

This proves that unless $W = 0$ probability is not conserved.

3.3 a) In deep water the phase velocity of water waves of wavelength λ is

$$v = \sqrt{\frac{g\lambda}{2\pi}}\quad.$$

What is the group velocity?
b) The phase velocity of a typical electromagnetic wave in a wave guide has the form

$$v = \frac{c}{\sqrt{1-(\omega_0/\omega)^2}}$$

where ω_0 is a certain characteristic frequency. What is the group velocity of such waves?

3.4 Which of the following operators are linear?

a)

$$K\psi(x) = \int K(x,y)\psi(y)dy \; .$$

b) K^3 where K is defined above.

c) AB if A and B are linear.

d) B^{-1} if B is linear and B^{-1} is defined by $B^{-1}B = BB^{-1} = 1$.

e) $\exp A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$ if A is linear.

f) $A\psi = \exp(\lambda\psi)$.

3.5 a) Compute, in closed form, the probability density $\rho(t,x)$ for the wave-function.

$$\Psi(t,x) = \int_{-\infty}^{\infty} dk\, A(k)\, \exp -i\left(\frac{\hbar k^2}{2m}t - kx\right)$$

where

$$A(k) = e^{-L^2k^2/2} \; .$$

b) What is the "width" of the probability density at time $t = 0$ and at time t ?

3.6 Show that if we write

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = \frac{\hbar}{i}\mathbf{r} \times \nabla$$

then

$$\begin{aligned} L_xL_y - L_yL_x &= i\hbar L_z \\ L_yL_z - L_zL_y &= i\hbar L_x \\ L_zL_x - L_xL_z &= i\hbar L_y \; . \end{aligned} \qquad (3.9.82)$$

These relations will be used later.

3.7 Assume a beam of free particles is moving along the x-axis with velocity v is such that there is one particle in a volume V.

a) What is the corresponding normalized, time-dependent wave function for such a particle?

b) What is the number of particles crossing a unit area, normal to the x-axis, per unit time?

3.8 A free particle has the wave packet at time $t = 0$ given by:

$$\Psi(0,x) = \int_{-\infty}^{\infty} dk \, \frac{e^{ikx}}{k^2 + a^2} \quad .$$

Determine an expression for the wavefunction for a later time t . Do not attempt to evaluate the resulting integral.
Hint: Use equation (3.5.18) and remember that the energy for a free particle is given by equation (3.5.16).

3.9 You are given the classical Hamiltonian for the motion of a particle in the form

$$H = \frac{p^2}{2m} e^{-\alpha x}$$

where α is a constant. Find an acceptable Hamiltonian operator for this system. Notice that the answer is by no means unique.

3.10 In classical mechanics, the reference level for the potential energy is arbitrary. What are the effects on the wavefunction and energy of adding a constant potential in the time-dependent Schrödinger equation?

Bibliography

[3.1] G.P. Thomson, Proc. Roy. Soc. **A117**, 600 (1928). E. Rupp. Ann. d. Physik **85**, 981 (1928).

[3.2] C. Jönsen, Z. Physik **161**, 454 (1961). A translation of this paper is given by D. Brandt and S. Hirschi, Am. J. Phys. **42**, 5 (1974).

[3.3] M. Born, Z. Physik **37**, 863 (1926). M. Born, Nature **119**, 354 (1972)

[3.4] For a treatment of the construction of quantum mechanical operators and the possible ambiguities see:
G.R. Gruber, Am. J. Phys. **40**, 1537-1538 (1972) .
G.R. Gruber, Foundations of Phys. **1**, 227-234 (1971).
G.R. Gruber, International Journal of Theoretical Physics **61**, 31-35 (1972).
also
J.R. Shewell, Amer. J. Phys. **27**, 16 (1959).

[3.5] H. Goldstein, *Classical Mechanics* - Addison Wesley Publishing Co., Reading, Mass., U.S.A. (1959) Sections 1-5.

Chapter 4

One-Dimensional Problems

4.1 Introduction

The purpose of this chapter is to further build up our familiarity with the Schrödinger equation by solving several problems. We begin by classifying two types of problems. Then we study a model of a particle in a box and look for the corresponding energy levels. This leads us naturally to consider the effect of mirror symmetry or parity. The next problem we treat, in considerable detail, is scattering from a potential with a step. We then examine the energy levels for a particle in a finite potential well. Next we study tunneling through a square potential barrier. Finally we briefly consider the concept of time reversal and how to include it in a quantum mechanical treatment.

Starting with the Schrödinger equation

$$H\Psi = i\hbar\frac{\partial\Psi}{\partial t} \tag{4.1.1}$$

we look for stationary state solutions. These are of the form

$$\Psi(t,x) = \psi_E(x)e^{-\frac{iE}{\hbar}t} \ . \tag{4.1.2}$$

Henceforth we suppress the subscript E in ψ_E, and simply write ψ. Then ψ satisfies the time independent Schrödinger equation

$$H\psi = E\psi \ . \tag{4.1.3}$$

For one dimensional systems we consider Hamiltonians of the form

$$H = T + V = \frac{p^2}{2m} + V(x) \ . \tag{4.1.4}$$

Thus, the Schrödinger equation (4.1.3) becomes:

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi = E\psi \ . \tag{4.1.5}$$

In order to develop our intuition we study this equation for some simple solvable cases. The problems to be considered can be divided into two cases.

1) Scattering: In this case $E \geq 0$. The potential may be attractive $V < 0$ or repulsive $V > 0$.

2) Bound States: In this case we generally have $E \leq 0$, except if the potential has infinitely high walls. If the potential does not have infinitely high walls we use our freedom in defining energy to within an arbitrary constant to choose the zero of energy to have $E \leq 0$. The potential must be attractive.

Before proceeding, we rewrite (4.1.5) in dimensionless form. Call

$$U(x) \equiv \frac{2m}{\hbar^2} V(x) \tag{4.1.6}$$

and

$$k^2 = \frac{2mE}{\hbar^2} \ . \tag{4.1.7}$$

Then we get

$$\frac{d^2\psi}{dx^2} + (k^2 - U)\psi = 0 \ . \tag{4.1.8}$$

This is the equation we shall study for a while. To specify our solutions further we state that we are looking for solutions on $-\infty < x < \infty$ such that they are

1. everywhere finite
2. continuous and differentiable.

Now as we saw before, (Problem 3.2) $V(x)$ has to be real in order to ensure conservation of probability. Thus, if our boundary conditions are real, we are looking for real solutions of (4.1.8). This is always the case for bound state problems. For scattering problems, the boundary conditions are complex since they specify some incoming wave of the form (3.5.18).

4.2 Particle in a Box

From a comparison of geometrical and wave optics we expect wave phenomena to be most prominent when the index of refraction changes rapidly. Thus, quantum effects are most prominent if $V(x)$ and hence $U(X)$ vary rapidly in one wavelength. This is always the case if $U(x)$ has a step. Thus, to see this effect and for simplicity, we consider the following simple potential

$$U(x) = \begin{cases} 0 & \text{if } |x| < a \\ \infty & \text{if } |x| > a \end{cases} \ . \tag{4.2.9}$$

This potential models a particle in a one-dimensional box $|x| < a$. The potential has impenetrable walls, in other words, it is infinitely high.

Thus, we have to solve

$$\frac{d^2\psi}{dx^2} + k^2\psi = 0 \tag{4.2.10}$$

subject to

$$\psi = 0 \text{ at } x = \pm a \ . \tag{4.2.11}$$

The boundary conditions (4.2.11) follow from the fact that the wave-function is continuous and that the particle is confined to the box so that there must be zero probability for finding the particle outside the box. The solutions are

$$\psi_{e,k}(x) = A_k \cos(kx) + B_k \sin(kx) \quad -a < x < a \ . \tag{4.2.12}$$

Applying the boundary conditions at $x = \pm a$

$$\psi_k(\pm a) = 0 \tag{4.2.13}$$

we obtain the two conditions

$$\begin{aligned} A_k \cos(ka) + B_k \sin(ka) &= 0 \\ A_k \cos(ka) - B_k \sin(ka) &= 0 \ . \end{aligned} \tag{4.2.14}$$

Thus we have two kinds of possible solutions

$$\begin{aligned} B_k &= 0 \quad \cos(ka) = 0 \\ A_k &= 0 \quad \sin(ka) = 0 \ . \end{aligned} \tag{4.2.15}$$

So we have two sets of eigenvalues:

$$\begin{aligned} B_k &= 0 \quad ka = (n + 1/2)\pi \quad n = 0, 1, 2, \ldots \\ A_k &= 0 \quad ka = n\pi \qquad n = 1, 2, 3, \ldots \ . \end{aligned} \tag{4.2.16}$$

In the first case the energy eigenvalues are

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 \pi^2}{8ma^2}(2n+1)^2 \quad n = 0, 1, 2, 3, \ldots \tag{4.2.17}$$

and the wavefunction is

$$\psi_n(x) = A_n \cos\left(\frac{(2n+1)\pi x}{2a}\right) \quad -a < x < a \ . \tag{4.2.18}$$

In the second case the energy eigenvalues are

$$E_n = \frac{\hbar^2 \pi^2}{8ma^2}(2n)^2 \quad n = 1, 2, 3, \cdots \tag{4.2.19}$$

and the wavefunction is

$$\psi_n(x) = B_n \sin\left(\frac{n\pi x}{a}\right) \quad -a < x < a \ . \tag{4.2.20}$$

The infinite square well is not as unphysical or artificial as it might appear at first. In fact, it is a good mathematical representation of a particle confined to a (one-dimensional) box. In such a box, a measurement of the energy of the particle yields one of the energies we computed. However, for macroscopic boxes, the energy levels are so closely spaced that they are experimentally indistinguishable from a continuum of energy levels.

4.3 Parity

In the previous example we found two classes of solutions corresponding to

Class 1. $\psi(-x) = \psi(x)$ and

Class 2. $\psi(-x) = -\psi(x)$.

This is a consequence of an interesting property of the Hamiltonian.

Consider a general Hamiltonian

$$H = \frac{p^2}{2m} + V(x) \tag{4.3.21}$$

where

$$V(-x) = V(x) \ . \tag{4.3.22}$$

The Schrödinger equation reads:

$$H\psi = E\psi \tag{4.3.23}$$

and due to the symmetry of $V(x)$ we find that whenever $\psi(x)$ is a solution then so is the even combination

$$\psi_+(x) = \frac{1}{\sqrt{2}}[\psi(x) + \psi(-x)] \tag{4.3.24}$$

as well as the odd combination

$$\psi_-(x) = \frac{1}{\sqrt{2}}[\psi(x) - \psi(-x)] \ . \tag{4.3.25}$$

Of course if ψ is an even function then ψ_- vanishes and if ψ is an odd function then ψ_+ vanishes.

These considerations lead us to introduce an operator P called the *parity operator* such that

$$(P\psi)(x) = \psi(-x) \ . \tag{4.3.26}$$

That is, P operating on a function ψ evaluated at x gives the same function with a negative argument $\psi(-x)$. In effect if a state is described by $\psi(x)$ then $(P\psi)(x)$ describes the mirror image of this state. We henceforth simplify the notation and write $(P\psi)(x)$ simply as $P\psi(x)$. The eigenvalue equation for P is particularly simple since if

$$P\psi(x) = \lambda\,\psi(x) \ . \tag{4.3.27}$$

Then,

$$P^2\psi(x) = \lambda^2\,\psi(x) \ . \tag{4.3.28}$$

But,

$$P^2\psi(x) = P\psi(-x) = \psi(x) \ . \tag{4.3.29}$$

So $\lambda^2 = 1$ and the eigenvalues of P are ± 1. Clearly the functions ψ_+ , ψ_- are eigenfunctions of P since

$$P\psi_+ = +1 \cdot \psi_+ \tag{4.3.30}$$

and

$$P\psi_- = -1 \cdot \psi_- \ . \tag{4.3.31}$$

If we now consider the Schrödinger equation with a potential V which is an even function of x so that $V(-x) = V(x)$ and apply P to both sides of the Schrödinger equation we find since

$$\frac{p^2}{2m} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \tag{4.3.32}$$

is also even under reflection that

$$PH\psi = PHP^2\psi = HP\psi = EP\psi \tag{4.3.33}$$

where we have used the fact that

$$P^2 = 1 \quad \text{and} \quad PV(x)P = V(-x)P^2 = V(x) \ . \tag{4.3.34}$$

Thus, we have that if ψ is an eigenfunction of H then so is $P\psi$. So we see that it is possible for ψ to be an eigenfunction of H and P simultaneously. In fact, quite generally, as we shall see later, if two observables commute (in this case $HP = PH$) one can solve the eigenvalue problem for them simultaneously.

In general, whenever the Hamiltonian has some special symmetry, this reflects itself in the wave-functions. It is of great assistance to recognize and utilize this symmetry from the start as in the case of parity above.

4.4 Scattering from a Step-Function Potential

We next consider a scattering problem. This time we choose a step function potential

$$U(x) = \begin{cases} U_1 & \text{if } x < 0 \\ U_2 & \text{if } x > 0 \end{cases} \ . \tag{4.4.35}$$

Case 1.

$$U_1 < k^2 < U_2 \ . \tag{4.4.36}$$

Classically, the problem is as follows. We have a particle corresponding to energy $\hbar^2 k^2/2m$ come from $-\infty$, hit the potential U_2 and reflect back. The particle can not enter the region $x > 0$ since this would yield a negative kinetic energy. Thus, we call the region $x > 0$ the *classically forbidden region.*

In the quantum mechanical situation we must solve the wave equation both in the classically allowed as well as in the classically forbidden region. We expect

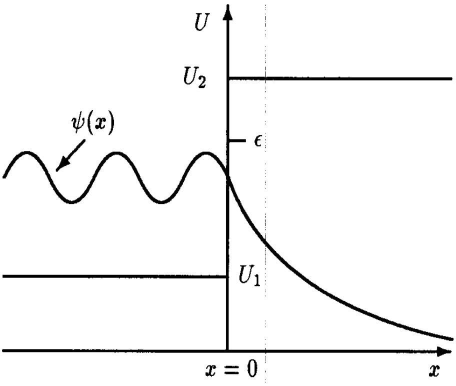

Figure 4.1: Scattering from a potential step: $V_1 < E < V_2$.

the wave to penetrate some distance into the forbidden region. Thus, we expect damped exponential behaviour (figure 4.1). The Schrödinger equation is:

$$\begin{aligned} \frac{d^2\psi}{dx^2} + k_1^2\psi &= 0 \quad x < 0 \\ \frac{d^2\psi}{dx^2} - k_2^2\psi &= 0 \quad x > 0 \ . \end{aligned} \tag{4.4.37}$$

Here,

$$\begin{aligned} k_1^2 &= k^2 - U_1 > 0 \\ k_2^2 &= U_2 - k^2 > 0 \ . \end{aligned} \tag{4.4.38}$$

We can therefore write the general solution as

$$\begin{aligned} \psi &= A\left[e^{ik_1x} + R\,e^{-ik_1x}\right] \quad x < 0 \\ \psi &= ASe^{-k_2x} \quad x > 0 \ . \end{aligned} \tag{4.4.39}$$

We have already imposed the condition that $\psi(x)$ must vanish for $x \to \infty$. Furthermore we have chosen a particular form for the solution for $x < 0$. Thus, the incoming wave $A\,e^{ik_1x}$ has been given an amplitude A.

4.4.1 Boundary Conditions

In most scattering problems in one dimension it is more convenient to match logarithmic derivatives

$$\frac{1}{\psi}\frac{d\psi}{dx}$$

rather than ψ and $d\psi/dx$ separately. This is because the normalization constants then cancel. In our case matching

$$\frac{1}{\psi}\frac{d\psi}{dx} \quad \text{at } x = 0$$

yields:

$$\frac{ik_1(1-R)}{1+R} = -k_2 \ . \tag{4.4.40}$$

This is an equation for R. So,

$$R = \frac{k_1 - ik_2}{k_1 + ik_2} \ . \tag{4.4.41}$$

Now, matching ψ at $x = 0$ gives

$$S = 1 + R = \frac{2k_1}{k_1 + ik_2} \ . \tag{4.4.42}$$

Notice that we still have one free parameter, say A . This corresponds to the fact that we are free to specify the flux of incoming particles, i.e. we can normalize ψ to yield the correct incoming flux. Before discussing the physics of this case we solve another case. We then discuss the physics for both these cases.
Case 2.

$$U_1 < U_2 < k^2 \ . \tag{4.4.43}$$

In this case we can have particles incident both from the left and right unlike Case 1 where particles were incident only from the left. Our theory, however, is linear and therefore we can consider the two cases (particles from the left and particles from the right) separately and superimpose the results. Physically this corresponds to the fact that particles coming from the left and from the right scatter independently from the potential. This means that we are omitting any possible interactions between the particles.

4.4.2 Particles from the Left

Suppose we normalize the flux coming from the left so that the incident wave is e^{ikx}. Then the solution of the Schrödinger equation under condition (4.4.43) is:

$$\begin{aligned} \psi &= e^{ik_1x} + Re^{-ik_1x} \quad x < 0 \\ \psi &= Se^{ik_2x} \quad x > 0 \ . \end{aligned} \tag{4.4.44}$$

In this case

$$\begin{aligned} k_1^2 &= k^2 - U_1 \\ k_2^2 &= k^2 - U_2 \ . \end{aligned} \tag{4.4.45}$$

Equating logarithmic derivatives at $x = 0$ gives:

$$\frac{ik_1 - ik_1R}{1+R} = ik_2 \ . \tag{4.4.46}$$

Hence

$$R = \frac{k_1 - k_2}{k_1 + k_2} \tag{4.4.47}$$

or

$$R = \frac{\sqrt{E - V_1} - \sqrt{E - V_2}}{\sqrt{E - V_1} + \sqrt{E - V_2}} \, . \tag{4.4.48}$$

Furthermore, equating the wave functions at $x = 0$ yields

$$S = 1 + R = \frac{2k_1}{k_1 + k_2} \, . \tag{4.4.49}$$

R gives the amplitude of the reflected wave, S gives the amplitude of the transmitted wave. We have to multiply $|S|^2$ by appropriate velocity factors to get the actual transmitted flux. This we do later.

If instead we had started with a beam of particles incident from the right instead of the left we would look for a solution of the form

$$\begin{aligned} \psi &= e^{-ik_2x} + Re^{ik_2x} \quad x > 0 \\ &= Se^{-ik_1x} \quad x < 0 \end{aligned} \tag{4.4.50}$$

and impose continuity of ψ and $d\psi/dx$ at $x = 0$ to solve for R and S.

We now turn to an examination of the physics involved in these solutions. To do this we rewrite the solutions slightly and compare them with the classical problem.

Case 1.

$$V_1 < E < V_2 \, . \tag{4.4.51}$$

As we have already stated, classically this corresponds to a particle of energy $E = (\hbar^2 k^2/2m)$ incident from the left with velocity $v = \hbar k_1/m$. At $x = 0$ it bounces elastically off the potential V_2 (i.e. it receives an impulse) and starts to travel back towards $x = -\infty$ with the original velocity v_1 .

We now examine the quantum mechanical case. The biggest difference between the classical and quantum mechanical description lies in the fact that classically the particle can never enter the region $x > 0$ where $V_2 > E$. In this case we would have

$$E = T + V_2 < V_2 \tag{4.4.52}$$

where T is the kinetic energy. This would mean

$$T = \frac{1}{2}mv^2 < 0 \tag{4.4.53}$$

and cannot occur. On the other hand, quantum mechanically we see that $\psi \neq 0$ for $x > 0$. In fact ψ is only damped exponentially. We now examine the second case.

Case 2. Both classically and quantum mechanically there are two possible motions in this case. The particles can be incident from the left or the right. We discuss only the former case.

Classically the particle is incident from the left with a velocity

$$v_1 = \frac{\hbar k_1}{m} \tag{4.4.54}$$

experiences a sudden change in velocity at $x = 0$ (an impulse) and continues to the right at a velocity

$$v_2 = \frac{\hbar k_2}{m} . \tag{4.4.55}$$

Quantum mechanically this motion is described as follows. We have

$$\psi = \begin{cases} e^{ik_1x} + R\,e^{-ik_1x} & x < 0 \\ S\,e^{ik_2x} & x > 0 \end{cases} \tag{4.4.56}$$

where, as we found earlier

$$R = \frac{k_1 - k_2}{k_1 + k_2} . \tag{4.4.57}$$

Thus, unlike the classical case, there is now a non-zero probability that the particle is reflected at the discontinuity in the potential. This is similar to light hitting a window pane, some is transmitted and some is reflected.

In order to get a better understanding of the two coefficients R and S we return to our concept of probability current. Recall that the probability density is given by

$$\rho = \Psi^*\Psi = |\Psi|^2 \tag{4.4.58}$$

and the probability density current is given by

$$j = \frac{-i\hbar}{2m}(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*) . \tag{4.4.59}$$

As we saw earlier (section 3.3) these quantities satisfy the conservation law

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0 , \tag{4.4.60}$$

the equation of continuity. It is this equation that puts us finally in a position to interpret the coefficients R and S.

4.4.3 Interpretation of R and S

We have the solution

$$\psi(x) = \begin{cases} \psi_< = e^{ik_1x} + R\,e^{-ik_1x} & \text{if } x < 0 \\ \psi_> = S\,e^{ik_2x} & \text{if } x > 0 \end{cases} \tag{4.4.61}$$

where

$$R = \frac{k_1 - k_2}{k_1 + k_2} \tag{4.4.62}$$

$$S = \frac{2k_1}{k_1 + k_2} \ . \tag{4.4.63}$$

Throughout this problem we have velocities v_i given by $v_i = (\hbar k_i)/m$. If we compute the probability density flux j for $x < 0$ we get

$$\begin{aligned} j(x) &= -\frac{i\hbar}{2m}\left(\psi^*_< \frac{\partial \psi_<}{\partial x} - \psi_< \frac{\partial \psi^*_<}{\partial x}\right) \\ &= -\frac{i\hbar}{2m}\left[\left(e^{-ik_1x} + R^*\, e^{ik_1x}\right) ik_1 \left(e^{ik_1x} - R\, e^{-ik_1x}\right)\right] \\ &\quad +\frac{i\hbar}{2m}\left[\left(e^{ik_1x} + R\, e^{-ik_1x}\right)(-ik_1)\left(e^{-ik_1x} - R^*\, e^{ik_1x}\right)\right] \\ &= \frac{\hbar k_1}{m}\left[1 - R\, e^{-i2k_1x} + R^*\, e^{i2k_1x} - |R|^2\right] \\ &\quad +\frac{\hbar k_1}{m}\left[1 + R\, e^{-i2k_1x} - R^*\, e^{i2k_1x} - |R|^2\right] \end{aligned} \tag{4.4.64}$$

or

$$j(x) = \frac{\hbar k_1}{m}[1 - |R|^2] = v_1 - v_1 |R|^2 \ . \tag{4.4.65}$$

Thus, for $x < 0$ we have a flux v_1 to the right and a flux $v_1|R|^2$ to the left. This has the nice interpretation that $|R|^2$ is the fraction of the flux reflected and hence $|R|^2$ is the reflection probability $\mathcal{R}$.

For $x > 0$ we get

$$\begin{aligned} j(x) = &-\frac{i\hbar}{2m}\left[S^*\, e^{-ik_2x}(ik_2) S\, e^{ik_2x}\right] \\ &+\frac{i\hbar}{2m}\left[S\, e^{ik_2x}(-ik_2) S^*\, e^{-ik_2x}\right] \end{aligned} \tag{4.4.66}$$

or

$$j(x) = \frac{\hbar k_2}{m}\, |S|^2 = v_2\, |S|^2 \ . \tag{4.4.67}$$

This can be rewritten as

$$j(x) = \frac{\hbar k_1}{m}\frac{k_2}{k_1}|S|^2 = v_1 \frac{k_2}{k_1}|S|^2 \ . \tag{4.4.68}$$

The fraction of the incident flux transmitted is therefore given by

$$\frac{v_1 \frac{k_2}{k_1}|S|^2}{v_1} = \frac{k_2}{k_1}|S|^2 \ . \tag{4.4.69}$$

Thus, the transmission probability $\mathcal{T}$ is

$$\mathcal{T} = \frac{k_2}{k_1}|S|^2 \ . \tag{4.4.70}$$

As a check we have that a particle is either transmitted past the potential step or reflected. Thus, we need to find that

$$\mathcal{T} + \mathcal{R} = 1 . \tag{4.4.71}$$

But this gives

$$\frac{k_2}{k_1}\left(\frac{2k_1}{k_1+k_2}\right)^2 + \left(\frac{k_1-k_2}{k_1+k_2}\right)^2 = \frac{4k_1k_2+(k_1-k_2)^2}{(k_1+k_2)^2} = 1 \tag{4.4.72}$$

as required.

Returning to Case 1 where $k^2 < U_2$ we find that the transmitted current for $x > 0$ is zero. On the other hand the reflection coefficient $\mathcal{R}$ is given by

$$\mathcal{R} = |R|^2 = \left|\frac{k_1 - ik_2}{k_1 + ik_2}\right|^2 = 1 . \tag{4.4.73}$$

This clearly shows that, even though the beam penetrates into the classically forbidden region, all of the beam is reflected, just like in the classical case.

4.5 Finite Square Well: Bound States

Unlike the case of an infinitely deep square well (a particle in a box) we now look at a finite potential well.

$$U(x) = \begin{cases} 0 & \text{if } |x| > a \\ -U & \text{if } |x| < a \end{cases} \tag{4.5.74}$$

where $U > 0$. For bound state problems we call

$$\kappa^2 = \frac{2m|E|}{\hbar^2} = -\frac{2mE}{\hbar^2} \tag{4.5.75}$$

since $E < 0$ and

$$\frac{2m(E+U)}{\hbar^2} = k^2 . \tag{4.5.76}$$

Thus, the Schrödinger equation (4.1.8) becomes

$$\begin{aligned} \frac{d^2\psi}{dx^2} - \kappa^2\psi = 0 \qquad |x| &> a \\ \frac{d^2\psi}{dx^2} + k^2\psi = 0 \qquad |x| &< a . \end{aligned} \tag{4.5.77}$$

Clearly the potential is symmetric under $x \to -x$. This means that the Hamiltonian does not change under parity and we can look for even and odd parity solutions. Furthermore we want $\int_{-\infty}^{\infty} |\psi|^2\, dx$ to represent the probability of finding the particle somewhere since $|\psi|^2$ is the probability density. So,

$$\int_{-\infty}^{\infty} |\psi|^2\, dx = 1 . \tag{4.5.78}$$

In order that (4.5.78) be possible it imposes the restriction that $\psi(x)$ vanishes as $x \to \pm\infty$. Thus, we get the following two classes of solutions.
Even parity

$$\psi_+(x) = \begin{cases} A\, e^{\kappa x} & x < -a \\ B\, \cos kx & |x| < a \\ A\, e^{-\kappa x} & x > a \end{cases} \quad . \tag{4.5.79}$$

Odd parity

$$\psi_-(x) = \begin{cases} A\, e^{\kappa x} & x < -a \\ B\, \sin kx & |x| < a \\ -A\, e^{-\kappa x} & x > a \end{cases} \quad . \tag{4.5.80}$$

From the differential equation (4.5.77) it follows by integrating about the points $x = \pm a$ that both ψ and $d\psi/dx$ are continuous at these points. Therefore, imposing continuity of the logarithmic derivative we obtain
Even parity

$$\kappa = k \tan ka \tag{4.5.81}$$

and

$$A = B\, e^{\kappa a}\, \cos ka \ . \tag{4.5.82}$$

Odd parity

$$\kappa = -k \cot ka \tag{4.5.83}$$

and

$$A = -B\, e^{\kappa a}\, \sin ka \ . \tag{4.5.84}$$

In either cases it requires numerical or graphical techniques to determine the energy eigenvalues. For example, setting

$$ka = x \quad \text{and} \quad \kappa a = y \tag{4.5.85}$$

we get

$$x^2 + y^2 = (k^2 + \kappa^2)a^2 = U a^2 \tag{4.5.86}$$

or

$$x^2 + y^2 = \frac{2ma^2}{\hbar^2} U \equiv R^2 \ . \tag{4.5.87}$$

In addition the two eigenvalue equations are:

Case 1:

$$x \tan x = y \tag{4.5.88}$$

and

Case 2:

$$x \cot x = -y \ . \tag{4.5.89}$$

It is now an easy matter to solve these equations graphically or numerically.

If $|x| << 1$ we can also get an approximate solution for Case 1 since then

$$x^2 \approx y \quad x^2 + y^2 = R^2 \ . \tag{4.5.90}$$

So,

$$y^2 + y - R^2 = 0 \tag{4.5.91}$$

and

$$y \approx -\frac{1}{2} \pm \frac{1}{2}\sqrt{1 + 4R^2} \ . \tag{4.5.92}$$

Only the + sign applies. Furthermore $R^2 < 1$ so we get

$$y \approx \frac{1}{2}[1 + 2R^2 - 1] \approx R^2 \ . \tag{4.5.93}$$

But

$$y = \kappa a = \sqrt{-\frac{2mE}{\hbar^2}}\, a \ . \tag{4.5.94}$$

Therefore,

$$E = -\frac{\hbar^2 R^4}{2ma^2} \tag{4.5.95}$$

and substituting back all the values

$$E = -\frac{2ma^2}{\hbar^2} U^2 \ . \tag{4.5.96}$$

This is an approximation for the lowest eigenvalue. For values of the numerical solution see the book by L. Schiff [4.1].

4.6 Tunneling Through a Square Barrier

Classically a particle is never found in a region in which $V > E$, but as we saw previously, in the quantum mechanical case there is a non-zero probability of finding a particle in a classically forbidden region although the wave function decreases exponentially the further we penetrate into the forbidden region. If the forbidden region is not too broad, the tail of the wave function can penetrate part of it and thus, it is possible for a quantum particle to penetrate a potential barrier. This is known as tunneling. To illustrate this, consider the following problem.

We have a potential

$$V(x) = \begin{cases} V_0 > 0 & 0 < x < a \\ 0 & x < 0 \text{ or } x > a \,. \end{cases} \tag{4.6.97}$$

A wave e^{ikx} (corresponding to a particle of definite momentum $\hbar k$) is incident from the left. We assume

$$E = \frac{\hbar^2 k^2}{2m} < V_0 \,. \tag{4.6.98}$$

Thus, the region $0 < x < a$ is classically forbidden and a classical particle incident from the left would simply rebound elastically in this case. In the quantum mechanical case we must solve the Schrödinger equation which can be written

$$\begin{aligned} \frac{d^2\psi}{dx^2} + k^2\psi &= 0 \quad x < 0 \text{ or } x > a \\ \frac{d^2\psi}{dx^2} - \kappa^2\psi &= 0 \quad 0 < x < a \end{aligned} \tag{4.6.99}$$

where

$$\kappa^2 = \frac{2m}{\hbar^2}(V_0 - E) \,. \tag{4.6.100}$$

The solution for $x < 0$ is dictated by the physical boundary condition that we have an incoming wave (whose amplitude we are free to choose and therefore set equal to 1) plus a reflected wave

$$\psi = e^{ikx} + R\,e^{-ikx} \quad x < 0 \,. \tag{4.6.101}$$

Also, for $x > a$ we have a purely transmitted wave travelling to the right. Here we have chosen the physical condition to correspond to no wave incident from the right.

$$\psi = S\,e^{ikx} \quad x > a \,. \tag{4.6.102}$$

The transmission coefficient $\mathcal{T}$ is given by

$$\mathcal{T} = \frac{k}{k}|S|^2 = |S|^2 \,. \tag{4.6.103}$$

Also, for $0 < x < a$ we must take the most general solution so that we have

$$\psi = A\,e^{\kappa x} + B\,e^{-\kappa x} \quad 0 < x < a\,. \tag{4.6.104}$$

We now match the wave function and its derivative at both $x = 0$ and $x = a$.
$x = 0$

$$\begin{aligned} 1 + R &= A + B && (4.6.105)\\ ik(1 - R) &= \kappa(A - B) && (4.6.106) \end{aligned}$$

$x = a$

$$\begin{aligned} S\,e^{ika} &= A\,e^{\kappa a} + B\,e^{-\kappa a} && (4.6.107)\\ ikS\,e^{ika} &= \kappa\left[A\,e^{\kappa a} + B\,e^{-\kappa a}\right] \;. && (4.6.108) \end{aligned}$$

We can solve these four equations in four unknowns. However, only R and S are physically interesting. We get:

$$\begin{aligned} 2ik &= A(\kappa + ik) - B(\kappa - ik) && (4.6.109)\\ 0 &= A(\kappa - ik)e^{\kappa a} - B((\kappa + ik)e^{-\kappa a} \;. && (4.6.110) \end{aligned}$$

Thus,

$$\begin{aligned} A &= \frac{2ik(\kappa + ik)}{(\kappa + ik)^2 - (\kappa - ik)^2 e^{2\kappa a}} && (4.6.111)\\ B &= \frac{2ik(\kappa - ik)e^{2\kappa a}}{(\kappa + ik)^2 - (\kappa - ik)^2 e^{2\kappa a}} \;. && (4.6.112) \end{aligned}$$

Now,

$$\begin{aligned} R &= \frac{1}{2ik}[-A(\kappa - ik) + B(\kappa + ik)]\\ S &= \frac{e^{-ika}}{2ik}[A(\kappa + ik)e^{\kappa a} + B(\kappa - ik)e^{-\kappa a}] \;. && (4.6.113) \end{aligned}$$

Substituting for A and B yields:

$$R = \frac{(\kappa^2 + k^2)(e^{\kappa a} - e^{-\kappa a})}{(\kappa + ik)^2 e^{-\kappa a} - (\kappa - ik)^2 e^{\kappa a}} or \tag{4.6.114}$$

or

$$R = \frac{(\kappa^2 + k^2)\sinh \kappa a}{(k^2 - \kappa^2)\sinh \kappa a + 2ik\kappa \cosh \kappa a} \;. \tag{4.6.115}$$

Similarly,

$$S = \frac{2ik\kappa e^{-ika}}{(k^2 - \kappa^2)\sinh \kappa a + 2ik\kappa \cosh \kappa a} \;. \tag{4.6.116}$$

The transmission coefficient $\mathcal{T}$ is therefore given by:

$$\mathcal{T} = |S|^2 = \frac{4k^2\kappa^2}{(k^2-\kappa^2)^2\sinh^2\kappa a + 4k^2\kappa^2\cosh^2\kappa a} \, . \tag{4.6.117}$$

using $\cosh^2\kappa a - \sinh^2\kappa a = 1$ we can rewrite this to read

$$\mathcal{T} = \frac{4k^2\kappa^2}{(k^2+\kappa^2)^2\sinh^2\kappa a + 4k^2\kappa^2} \, . \tag{4.6.118}$$

Expressed in terms of the energy this reads

$$\mathcal{T} = \left[1 + \frac{V_0^2}{4E(V_0-E)}\sinh^2\kappa a\right]^{-1} \tag{4.6.119}$$

where as before

$$\kappa^2 a^2 = \frac{2ma^2}{\hbar^2}(V_0 - E) \, . \tag{4.6.120}$$

Thus, the transmission coefficient decreases exponentially as the width of the barrier increases. Also, as $E \to \infty$, $\mathcal{T} \to 1$ so that we get complete transmission. Again, as a check we have

$$\mathcal{T} + \mathcal{R} = 1 \tag{4.6.121}$$

where

$$\mathcal{R} = |R|^2 \, . \tag{4.6.122}$$

Thus, the beam is either reflected or transmitted.

4.6.1 Resonance Transmission

The derivation above remains valid even if $E > V_0$. In this case the only modification required is that we replace $V_0 - E$ by $E - V_0$ so that κ becames imaginary and we replace it by iK. This means that the transmission coefficient becomes

$$\mathcal{T} = \left[1 + \frac{V_0^2}{4E(E-V_0)}\sin^2 Ka\right]^{-1} \, . \tag{4.6.123}$$

It now follows that if $Ka = n\pi \quad n = 1, 2, \ldots$ then

$$\mathcal{T} = 1 \tag{4.6.124}$$

so that all of the beam is transmitted. This is known as *resonance transmission.* Notice that the resonance condition, $Ka = n\pi$, when written out gives energies

$$E_n = V_0 + \frac{\hbar^2\pi^2}{2ma^2}n^2 \tag{4.6.125}$$

which, except for the additive constant V_0 are the same as the energies for a particle in a box of width a.

The physics behind all this is very nice and we now discuss this. The complete reflected wave is a superposition of the wave reflected at $x = 0$ and the wave reflected at $x = a$. The optical path difference is therefore $2Ka$. This does not yet give the total phase difference between the two waves since the wave relected at $x = a$ is reflected at an optically more dense medium [1] and acquires an additional phase change of π. Thus, the total phase difference between the two waves is $2Ka + \pi$. To have no reflected wave, these two waves must interfere destructively this requires that the phase difference be an odd multiple of π. Writing this out we get

$$2Ka + \pi = (2n+1)\pi \; . \tag{4.6.126}$$

But this is precisely the condition we obtained above for resonance transmission.

4.7 Time Reversal

In the previous sections we found that the use of mirror symmetry, or parity, greatly facilitated the computation. There is another discrete symmetry which is frequently useful, it is called *time reversal*. Imagine a system such as perfectly elastic billiard balls on a frictionless surface. In viewing a movie of the collisions of these billiard balls we are unable to tell whether the movie is running forwards or backwards; the motions obtained by reversing the direction of the film or time, are equally physically possible. We state this by saying that the system is invariant under time reversal. To examine this a little more closely we first consider the following classical transformation T which maps t into $-t$. So,

$$Tt = -t \; . \tag{4.7.127}$$

From this it then follows that if the motion is reversible then

$$T\mathbf{x}(t) = \mathbf{x}(-t) = \mathbf{x}(t) \tag{4.7.128}$$

$$T\mathbf{p}(t) = \mathbf{p}(-t) = -\mathbf{p}(t) \tag{4.7.129}$$

and

$$TV(\mathbf{x}) = V(\mathbf{x}) \; . \tag{4.7.130}$$

Thus, if we have a system with a Hamiltonian

$$H = \frac{1}{2m}\mathbf{p}^2 + V(\mathbf{x}) \; , \tag{4.7.131}$$

and we change t to $-t$ then we obtain, as equations (4.7.128) to (4.7.130) show, the same Hamiltonian. This means (classically) that if a possible motion of the system (solution of the equations of motion) is given by $(\mathbf{x}(t), \mathbf{p}(t))$ then another

---

[1] Medium A is optically more dense than medium B if the wavelength in medium A is less than the wavelength in medium B

possible motion is given by $(\mathbf{x}(-t), \mathbf{p}(-t))$. This is one way to sometimes obtain new solutions from old ones.

We now look at this from the point of view of quantum mechanics. The Hamiltonian (4.7.131) which is invariant under the transformation (4.7.127), is our starting point. The time-dependent Schrödinger equation reads:

$$i\hbar\frac{\partial\Psi(t,x)}{\partial t} = H\Psi(t,x) \ . \tag{4.7.132}$$

If we complex conjugate this equation and then do the transformation

$$t \to t' = -t \tag{4.7.133}$$

we find

$$i\hbar\frac{\partial\Psi^*(-t',x)}{\partial t'} = H\Psi^*(-t',x) \tag{4.7.134}$$

where we have used the fact that H is invariant under $t \to -t$. Dropping the primes we see that $\Psi^*(-t,x)$ satisfies the same Schrödinger equation as $\Psi(t,x)$. We therefore define the action of the time reversal operator T on $\Psi(t,x)$ by

$$(T\Psi)(t,x) = \Psi^*(-t,x) \tag{4.7.135}$$

and $T\Psi$ is a solution of (4.7.132) if Ψ is a solution. This operator is anti-linear due to the complex conjugation. Thus, if λ is a complex number we find

$$T(\lambda\Psi) = \lambda^* T\Psi \ . \tag{4.7.136}$$

The operator T acting on other operators such as x and p has the following property

$$Tx = xT \tag{4.7.137}$$

$$Tp = -pT \tag{4.7.138}$$

and for the Hamiltonian (4.7.130), as we saw

$$TH = HT \ . \tag{4.7.139}$$

This particular way of implementing time-reversal is called Wigner [4.2] time-reversal.

4.8 Problems

4.1 An electron has an energy of 10 eV. For which of the following potentials will a classical approximation be valid?

a) A step function of height 10 eV, 0.1 eV, 10^{-3} eV.

b) A potential $V = V_0 \exp[-x^2/a^2]$ with $V_0 = 100$ eV $a = 10^{-15}$ cm, 10^{-10} cm, 10^{-5} cm, 1.0 cm.

4.2 A beam of particles of mass m moving from left to right encounters a sharp potential drop of amount V_0. Let E be the kinetic energy of the incoming particles and show that the fraction of particles reflected at the edge of the potential (located at $x = 0$) is given by

$$\left(\frac{\sqrt{E+V_0}-\sqrt{E}}{\sqrt{E+V_0}+\sqrt{E}}\right)^2 .$$

In view of this result, what will happen to a car moving at 10 km/hr if it meets the edge of a 200 m cliff? Is this answer reasonable? If not, why not?

4.3 A particle moving in one dimension interacts with a potential of the form

$$V(x) = 0 \qquad |x| > a$$

$$V(x) = -\frac{v_0}{2a} \quad |x| < a \ .$$

Find the equation determining the energy eigenvalues of this system. Solve it approximately assuming a is very small. What happens in the limit $a \to 0$?

4.4 The wavefunction of an electron in the ground state of a hydrogen-like atom is

$$\psi(r) = Ae^{-Zr/a} \ , \quad a = \frac{\hbar^2}{me^2}$$

where Z is the charge on the nucleus.
a) Determine the constant A, so that the wave- function is normalized to unity. Remember that you are in three dimensions and r represents the radial variable in spherical coordinates.
b) At what distance from the origin is the probability of finding the electron a maximum?
c) Determine the average value of: the kinetic energy, the potential energy and the total energy.
This verifies the virial theorem for the Coulomb potential.

4.5 Show that the wavefunction for a particle in a bound state may always be chosen to be real. By computing the current density give an explanation of the physical meaning of this result.

4.6 We found that for a particle in a box with sides at $x = \pm a$ the eigenfunctions and eigenvalues were:

$$\psi_{+,n} = A_n \cos[(n+1/2)\frac{\pi x}{a}] \ , \ E_{+,n} = (n+1/2)^2\frac{\hbar^2\pi^2}{2ma^2}$$

$$\psi_{-,n} = B_n \sin[n\frac{\pi x}{a}] \ , \ E_{-,n} = n^2 \frac{\hbar^2 \pi^2}{2ma^2} \ .$$

If we have a particle in such a box and its wavefunction at time $t = 0$ is given by

$$\Psi(0, x) = \frac{1}{\sqrt{a}} \sin[5\frac{\pi x}{a}]$$

find $\Psi(t, x)$.

4.7 Find the energy levels and wavefunctions for a particle in the potential

$$V(x) = \begin{cases} 0 & \text{if } x < 0 \, , \, x > a \\ V_0 & \text{if } 0 < x < a \end{cases} \ .$$

4.8 If a particle, in a box with sides at $x = \pm a$, is in a state described by the wavefunction

$$f(x) = A\cos(\pi x/2a) + B\sin(\pi x/a)$$

a) Choose A and B so that the particle is in the lowest possible energy state.
b) Choose A and B so that the state has parity $= +1$.

4.9 A beam of free particles with intensity N particles per second and energy E is incident on a potential. Find the number of particles that are reflected in one second. How many particles per second are transmitted?

4.10 A beam of free particles with intensity N_L particles per second and energy E_L is incident from the left on a potential

$$V(x) = \begin{cases} 0 & \text{if } 0 < x < a \\ \infty & \text{if } x < 0 \, , \, x > a \end{cases} \ .$$

At the same time another beam of free particles with intensity N_R particles per second and energy E_R is incident from the right. Calculate the total particle current travelling to the right.
Hint: The waves from the left and right are completely independent and their scattering from the potential can be handled independently.

4.11 The *Ramsauer-Townsend Effect*
Consider scattering of a particle with energy $E > V_0 > 0$ from a potential

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

Show that if the wavelength of the particle in the region $0 < x < a$ is such that $n\lambda = a$ then no reflection occurs.

The actual Ramsauer-Townsend Effect was observed in the scattering of electrons off atoms of the noble gases since due to their closed shell configurations these atoms have very sharp outer boundaries.

4.12 For the one-dimensional Schrödinger equation with potential $V(x)$ and any two independent solutions $u(x)$ and $v(x)$, corresponding to the same energy, show that the Wronskian

$$W(x) = u'(x)v(x) - u(x)v'(x)$$

is a constant.
Hint: Write out the Schrödinger equation for the two solutions and multiply each of the equations by the other solution. Use this to obtain a differential equation for W.

4.13 Consider a repulsive potential $V(x) > 0$ such that it vanishes outside a finite region $a < x < b$. Show that the reflection coefficient is the same for a particle incident from either the left or the right. Assume that the solutions in the interval $a < x < b$ have been normalized such that their Wronskian (see problem 4.11) is 1.

Bibliography

[4.1] L.I. Schiff *Quantum Mechanics*, McGraw-Hill Book Co., 2nd edition (1955) pages 36-40.
R.H. Dicke and J.P. Wittke, *Introduction to Quantum Mechanics* - Addison Wesley Publishing Co., Inc., Reading, Mass., U.S.A. (1960).
Many "modern" problems are formulated as one-dimensional problems and solved in H.J. Lipkin, *Quantum Mechanics* - North Holland Publishing Co., Amsterdam and London (1973).

[4.2] E.P. Wigner, *Group Theory* - Academic Press (1959) Chapter 26.

Chapter 5

More One-Dimensional Problems

5.1 Introduction

We continue our familiarization with the Schrödinger equation in this chapter. To understand how boundary conditions lead to discrete eigenvalues we begin with a qualitative discussion of the bound state problem. We then consider one of the most important solvable problems in both classical and quantum physics, the simple harmonic oscillator. We introduce this problem and solve it, for the first of several times, at this stage. In preparation for later needs we also introduce the concept of the delta function. To gain familiarity with it we solve two problems with it. These serve to illustrate that delta-function potentials are really a form of boundary condition. Finally we again consider scattering from a potential well and construct parity invariant solutions. This allows us to introduce the concept of phase shifts.

5.2 General Considerations

Before proceeding with a study of several more specific cases we pause and consider the one-dimensional bound state problem quite generally. The procedure will be heuristic but will hopefully give some insight. As always we start with the Schrödinger equation.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V\psi = E\psi \tag{5.2.1}$$

and rewrite it in the now familiar, canonical form

$$\frac{d^2\psi}{dx^2} - (\kappa^2 + U)\psi = 0 \tag{5.2.2}$$

since we are interested in the case where U is a potential well $U < 0$ and $-\kappa^2 = (2mE)/\hbar^2 < 0$. Furthermore, ψ is chosen to be real since we have real boundary conditions ($\psi \to 0$ as $|x| \to \infty$). Another property of the potential is that it is short range. This means $U \to 0$ for $|x| \to \infty$. Thus, for large x we have:

$$\frac{d^2\psi}{dx^2} = \kappa^2\psi \tag{5.2.3}$$

so that

$$\psi(x) \approx e^{-\kappa x} \quad \text{for } x \text{ large} \tag{5.2.4}$$

$$\psi(x) \approx e^{\kappa x} \quad \text{for } -x \text{ large.} \tag{5.2.5}$$

It is precisely these condition that lead to the quantization of energy levels as we now show. In fact the condition that ψ has to be square integrable is what makes this an eigenvalue problem and gives a discrete set of eigenvalues. Consider again equation (5.2.2)

$$\frac{d^2\psi}{dx^2} = (\kappa^2 + U)\psi\ . \tag{5.2.6}$$

and suppose we are in a classically permitted region so that $U - \kappa^2 \equiv -k^2 < 0$. Then,

$$\frac{d^2\psi}{dx^2} = -k^2\psi \tag{5.2.7}$$

in this region. Furthermore, if $\psi > 0$ then (5.2.7) states that $d^2\psi/dx^2 < 0$ so that ψ curves downwards as shown in figure 5.1. On the other hand, if $\psi < 0$ then $d^2\psi/dx^2 > 0$ and ψ curves up as shown in figure 5.2. Thus, in the classically permitted region, ψ always bends towards the x-axis and the behaviour may be described as oscillatory.

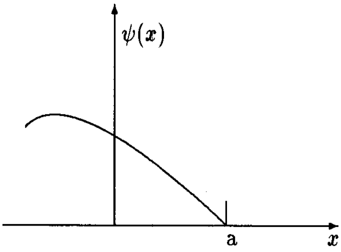

Figure 5.1: The classically permitted region with $\psi > 0$.

Suppose the classical bounds are given by $x = -a$, $x = b$. Thus, $U = \kappa^2$ at $x = -a$ and $x = b$ as shown in figure 5.3. Then, $-a < x < b$ is the classically

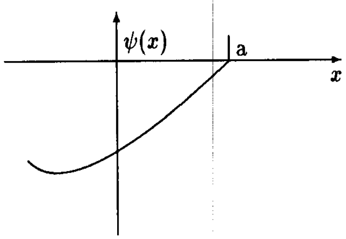

Figure 5.2: The classically permitted region with $\psi < 0$.

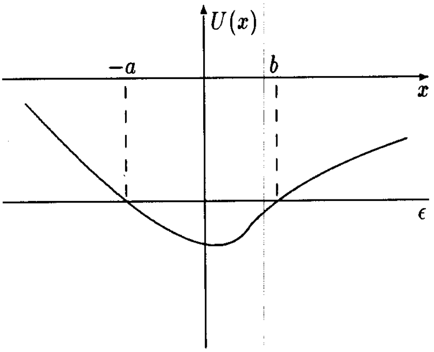

Figure 5.3: Classical turning points.

permitted region and in this region ψ is "oscillatory". In the classically forbidden regions ($x < -a$ and $x > b$) we have

$$\frac{d^2\psi}{dx^2} = k^2\psi \tag{5.2.8}$$

and ψ curves away from the axis. Now consider a solution as shown in figure 5.4. The solution is oscillatory between $-a < x < b$ and curves away from the x-axis for $b < x < -a$, but goes to $-\infty$ as $|x|$ increases. This solution has exponential growth rather than decay and is not square integrable and hence not physically acceptable. If k^2 is a little smaller, then ψ oscillates a little more slowly in the classically allowed region $-a < x < b$. In this case we may get a solution of the form shown in figure 5.5. Again the growth for large $|x|$ is exponential and the solution is unacceptable. On the other hand, it is possible to choose a value of k^2 between the two values just considered so that ψ behaves as shown in figure 5.6. In this case for large $|x|$ the wavefunction ψ is exponentially damped. Note that the value of k^2 so chosen means that the energy E is fixed. Thus, for a given number of nodes in the classically permitted region we get a fixed value

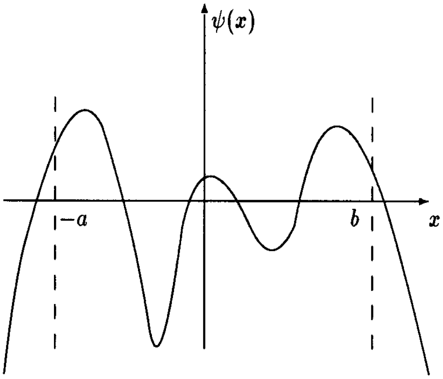

Figure 5.4: Solution for an energy that is slightly too large.

of the energy E and hence quantization.

Next we study several problems which besides being soluble are of intrinsic physical interest. The first of these is the simple harmonic oscillator. This is a problem that crops up over and over again in many different contexts.

5.3 The Simple Harmonic Oscillator

The Hamiltonian for this problem is

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 \, . \tag{5.3.9}$$

This gives us the time dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = H\Psi \, . \tag{5.3.10}$$

or

$$i\hbar\frac{\partial\Psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2\right)\Psi \, . \tag{5.3.11}$$

For the stationary states we put, as usual

$$\Psi(t,x) = e^{-iEt/\hbar}\psi(x) \tag{5.3.12}$$

and then the time independent Schrödinger equation reads

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2\right)\psi = E\psi \, . \tag{5.3.13}$$

As in the case of the square well we introduce dimensionless variables. To this end we set $\omega = \sqrt{k/m} =$ the classical frequency. Also, we set

$$\mu = \frac{2E}{\hbar\omega} \, , \tag{5.3.14}$$

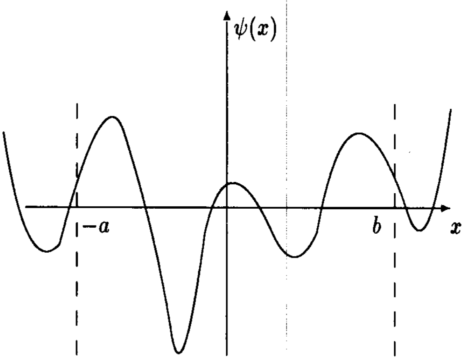

Figure 5.5: Solution for an energy that is slightly too small.

$$y = \left(\frac{mk}{\hbar^2}\right)^{1/4} x = \sqrt{\frac{m\omega}{\hbar}}\, x \,, \tag{5.3.15}$$

and

$$\psi(x) = u(y) \ . \tag{5.3.16}$$

Then, the time independent Schrödinger equation (5.3.13) becomes

$$\frac{d^2\psi}{dy^2} + (\mu - y^2)u = 0 \ . \tag{5.3.17}$$

Since we are looking for bound states the solutions must be damped for $y \to \pm\infty$. To ensure this we therefore first look at the asymptotic behaviour of u. For $|y|$ large we can neglect μ. Then,

$$\frac{d^2u}{dy^2} \approx y^2 u \ . \tag{5.3.18}$$

This yields that

$$u \approx e^{\pm 1/2\, y^2} \ . \tag{5.3.19}$$

Obviously, for a square-integrable u, we must choose the asymptotic solution with the minus sign. To solve the problem we now use the polynomial method of Sommerfeld [5.2]. Thus, we set

$$u = H(y)\, e^{-1/2\, y^2} \tag{5.3.20}$$

and look for the equation satified by $H(y)$. When (5.3.20) is inserted into (5.3.17) the resultant equation is

$$\frac{d^2H}{dy^2} - 2y\frac{dH}{dy} + (\mu - 1)H = 0 \ . \tag{5.3.21}$$

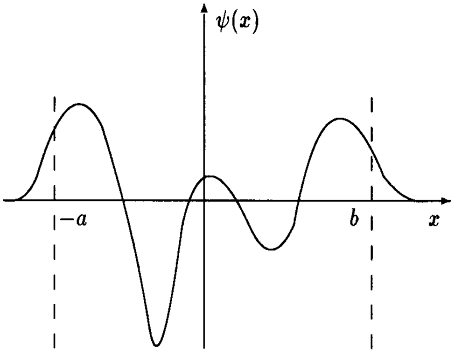

Figure 5.6: Bound state solution.

We solve this equation using the method of Frobenius by expanding the function H in a power series

$$H(y) = \sum_{n=0} a_n y^n \ . \tag{5.3.22}$$

Substituting this expansion into (5.3.21) gives:

$$\sum_{n=0} [n(n-1)a_n y^{n-2} - 2na_n y^n + (\mu - 1)a_n y^n] = 0 \ . \tag{5.3.23}$$

For $n \neq 0, 1$ we can replace n by $n' + 2$ in the first term to get

$$\sum_{n'=0} [(n'+2)(n'+1)a_{n'+2} - 2n'a_{n'} + (\mu - 1)a_{n'}]y^{n'} = 0 \ . \tag{5.3.24}$$

Hence, since the different powers of y are linearly independent, we get

$$a_{n+2} = \frac{2n+1-\mu}{(n+2)(n+1)} a_n \quad n = 0, 1, 2, 3, \dots \ . \tag{5.3.25}$$

This allows us to express
a_2 in terms of a_0 ,
a_3 in terms of a_1 ,
a_4 in terms of a_2 ,
a_5 in terms of a_3 , and so forth.
The two constants a_0 and a_1 are arbitrary. Furthermore, the Hamiltonian is invariant under the parity transformation that we introduced in section 4.3. Just as the parity operator was very useful for the square well considered in that section it is also very useful here since we find that it commutes with the Hamiltonian.

$$PH = HP \ . \tag{5.3.26}$$

Therefore, the eigenfunctions of the Hamiltonian H may be chosen to be either even or odd functions. For $a_0 = 0$, a_1 arbitrary we get odd solutions. For $a_1 = 0$, a_0 arbitrary we get even solutions. Thus, these are solutions of definite parity.

If we now consider the recursion relation (5.3.25) for large n we find that

$$a_{n+2} \approx \frac{2}{n+2} a_n \ . \tag{5.3.27}$$

However, if we expand

$$e^{y^2} = \sum \frac{1}{n!} y^{2n} \ , \tag{5.3.28}$$

the coefficient b_n of y^n is

$$b_n = \frac{1}{(n/2)!} \tag{5.3.29}$$

and

$$b_{n+2} = \frac{1}{(n/2+1)!} = \frac{1}{(n/2+1)} b_n = \frac{2}{n+2} b_n \ . \tag{5.3.30}$$

Thus, unless the series (5.3.22) terminates, $H(y)$ behaves asymptotically as e^{y^2} so that u behaves like $e^{y^2} e^{-y^2/2} = e^{y^2/2}$. This behaviour is unacceptable. Thus the series (5.3.22) must terminate. Note that this is the condition for u to be square integrable and leads to quantization. For the series to terminate requires that one of the a_n vanish. This will automatically happen if μ is an odd integer

$$\mu = 2N + 1 \ . \tag{5.3.31}$$

Then, $a_{N+2}, a_{N+4}, a_{N+6}$, etc. vanish. Thus, we have the quantization condition. Substituting back for μ we get the quantized energy levels

$$E_N = (N + 1/2)\hbar\omega \ . \tag{5.3.32}$$

Except for the additional $1/2\ \hbar\omega$ this is Planck's original assumption for the elementary oscillators of the electromagnetic field. There is a simple way to generate all these polynomials satisfying equation (5.3.21). With their normalization chosen in this way they are called hermite polynomials.

5.3.1 Generating Function for Hermite Polynomials

The function

$$S(t, y) = e^{-t^2+2ty} \tag{5.3.33}$$

is called the generating function for the hermite polynomials because as we now show

$$S(t, y) = e^{-t^2+2ty} = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(y) \ . \tag{5.3.34}$$

To see this we simply differentiate both sides of this equation and show that the functions H_n defined by this relation have to satisfy (5.3.21), the equation for the hermite polynomials with $\mu = 2n + 1$, where n is an integer. Thus,

$$\begin{aligned} \frac{\partial S}{\partial y} &= 2t\, e^{-t^2+2ty} = \sum_{n=0}^{\infty} \frac{2t^{n+1}}{n!} H_n(y) \\ &= \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n'(y) \ . \end{aligned} \tag{5.3.35}$$

Equating equal powers of t we find

$$H_n'(y) = 2nH_{n-1}(y) \ . \tag{5.3.36}$$

Also,

$$\begin{aligned} \frac{\partial S}{\partial t} &= (-2t + 2y)\, e^{-t^2+2ty} = \sum_{n=0}^{\infty} \frac{(-2t + 2y)t^n}{n!} H_n(y) \\ &= \sum_{n=0}^{\infty} \frac{t^{n-1}}{(n-1)!} H_n(y) \ . \end{aligned} \tag{5.3.37}$$

Again equating equal powers of t we get

$$H_{n+1}(y) = 2yH_n(y) - 2nH_{n-1}(y) \ . \tag{5.3.38}$$

The simplest (lowest order) differential equation involving only $H_n(y)$ that can be obtained from (5.3.36) and (5.3.38) is precisely (5.3.21). Thus, the functions $H_n(y)$ defined by the generating function (5.3.33) are precisely the hermite polynomials. The first five are listed below.

$$\begin{aligned} H_0(y) &= 1 \\ H_1(y) &= 2y \\ H_2(y) &= 4y^2 - 2 \\ H_3(y) &= 8y^3 - 12y \\ H_4(y) &= 16y^4 - 48y^2 + 12 \ . \end{aligned}$$

5.3.2 Rodrigues Formula for Hermite Polynomials

We again start with the generating function and realize that it may also be written as

$$S(t, y) = e^{-t^2+2ty} = e^{y^2} e^{-(t-y)^2} \ . \tag{5.3.39}$$

It therefore follows that

$$\frac{\partial S}{\partial t} = e^{y^2} \frac{\partial}{\partial t} e^{-(t-y)^2} = -e^{y^2} \frac{\partial}{\partial y} e^{-(t-y)^2} \ . \tag{5.3.40}$$

Repeating this n times we get

$$\frac{\partial^n S}{\partial t^n} = e^{y^2}\frac{\partial^n}{\partial t^n}e^{-(t-y)^2} = (-1)^n e^{y^2}\frac{\partial^n}{\partial y^n}e^{-(t-y)^2} \ . \tag{5.3.41}$$

This may be rewritten to read

$$\sum_m \frac{H_m(y)}{(m-n)!}t^{m-n} = (-1)^n e^{y^2}\frac{\partial^n}{\partial y^n}e^{-(t-y)^2} \ . \tag{5.3.42}$$

If we now set $t = 0$ the only term that survives in the sum on the left hand side is the term with $m = n$. Thus, we get

$$H_n(y) = (-1)^n e^{y^2}\frac{d^n}{dy^n}e^{-y^2} \ . \tag{5.3.43}$$

This is a Rodrigues formula for the hermite polynomials and is extremely useful.

5.3.3 Normalization

To normalize the wavefunctions

$$u_n(y) = A_n\, H_n(y)\, e^{-y^2/2} \tag{5.3.44}$$

we can use the generating function. Thus, we want to evaluate

$$\int_{-\infty}^{\infty} H_n(y)H_m(y)\, e^{-y^2}\, dy \ .$$

To this end we multiply the generating function with itself and integrate to get

$$\begin{aligned} &\int_{-\infty}^{\infty} e^{-t^2+2ty}\, e^{-s^2+2sy}\, e^{-y^2}\, dy \\ = \quad &\sum_{n,m=0}^{\infty} \frac{t^n s^m}{n!m!}\int_{-\infty}^{\infty} H_n(y)H_m(y)\, e^{-y^2}\, dy \ . \end{aligned} \tag{5.3.45}$$

The integrand on the left hand side may be rewritten to read

$$e^{-(t+s-y)^2}\, e^{2ts} \ .$$

So, this integral is just a Gaussian integral and yields

$$\sqrt{\pi}\, e^{2ts} = \sqrt{\pi}\sum_{n=0}^{\infty}\frac{(2ts)^n}{n!} \ . \tag{5.3.46}$$

Therefore, equating powers of t and s in the sums we have

$$\int_{-\infty}^{\infty} H_n(y)H_m(y)\, e^{-y^2}\, dy = \sqrt{\pi}\, 2^n\, n!\, \delta_{n,m} \ . \tag{5.3.47}$$

So the normalized eigenfunctions are

$$u_n(y) = \frac{1}{\pi^{1/4}}\frac{1}{\sqrt{2^n n!}}\, H_n(y)\, e^{-y^2/2} \ . \tag{5.3.48}$$

In terms of the unscaled variable x this becomes

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{\sqrt{2^n n!}} H_n(\sqrt{m\omega/\hbar}\, x)\, e^{-(m\omega/\hbar)\, x^2/2} . \tag{5.3.49}$$

We do not pursue the properties of the hermite polynomials any further since in section 9.1 we again solve the Schrödinger equation for the harmonic oscillator problem in a completely different way and there the normalization of the wave-function as well as other properties can be obtained in a much easier fashion.

Why is the simple harmonic oscillator (S.H.O.) so important? Aside from the fact that it can be solved in closed form, it provides a possible means of quantizing fields, for example, the electromagnetic. Furthermore, as we shall see later, the hermite functions $H_n(y)\exp(-y^2/2)$ are particularly convenient to work with. As a consequence they crop up in a large number of "practical" applications.

Another use of the S.H.O. is for estimating the ground state (lowest) energy of a system. This works well if the potential is smooth. Consider the potential shown (figure 5.7). Near the bottom, the potential is closely approximated by a parabola. This is in fact what a S.H.O. is. Thus, in the case shown the ground state energy is closely approximated by $1/2\, \hbar\omega - V_0$. To get the "spring constant" k we simply find the point $x = a$ where $V(x)$ is a minimum and evaluate $k = d^2V/dx^2|_{x=a}$. This is an extremely simple method for obtaining an estimate of the energy for low lying levels in the case of a complicated but smooth potential.

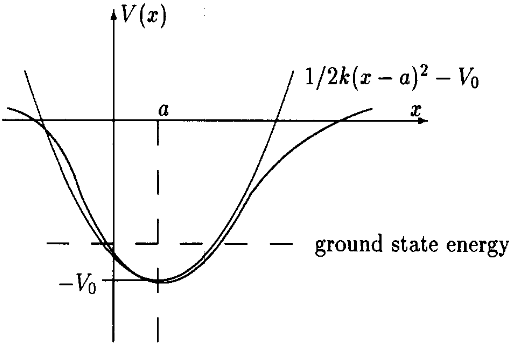

Figure 5.7: Simple Harmonic oscillator approximation for the ground state.

5.4 The Delta Function

There is one more class of potentials that occurs fairly often in practice due to the extreme simplicity of the solutions of the corresponding Schrödinger equation, the so-called delta function. This function plays a large role in later discussions and we therefore introduce it at this stage. Properly speaking, it is not a function at all and we define it rigorously in Chapter 7. For the time being, however, the following definitions suffice.

Definition

$$\delta(x) = 0 \quad x \neq 0$$

$\delta(x)$ is so singular at $x = 0$ that

$$\int_a^b f(x)\delta(x)\,dx = f(0) \tag{5.4.50}$$

whenever the interval (a, b) contains the origin.

The following integral representation is also valid for $\delta(x)$

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\,dk \tag{5.4.51}$$

where the integral is to be understood as

$$\delta(x) = \lim_{\epsilon\to 0+}\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx-\epsilon k^2}\,dk \tag{5.4.52}$$

and the limit is to be taken only after the delta function has itself been integrated with a smooth function. Thus, for example

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\delta(x)\,dx &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{ikx}f(x)\,dx\,dk \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{f}(k)\,dk \\ &= f(0)\end{aligned} \tag{5.4.53}$$

where

$$\tilde{f}(k) = \int_{-\infty}^{\infty} f(x)e^{ikx}\,dx \tag{5.4.54}$$

is the Fourier transform of $f(x)$. The inverse transform is given by

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{f}(k)\,e^{-ikx}\,dk \tag{5.4.55}$$

so that clearly

$$f(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{f}(k)\,dk\ . \tag{5.4.56}$$

Again, a rigorous proof of all these relations is delayed until Chapter 7.

Some of the more commonly used properties of the delta function (see problems 5.8 and 5.9) are listed below.

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)\,dx = f(a) \tag{5.4.57}$$

$$\delta(ax) = \frac{1}{a}\delta(x) \tag{5.4.58}$$

$$f(x)\delta(x-a) = f(a)\delta(x-a) \tag{5.4.59}$$

$$\int_{-\infty}^{\infty} \delta(x-y)\delta(y-z)\,dy = \delta(x-z) \tag{5.4.60}$$

$$\delta(x^2-a^2) = \frac{1}{2|a|}[\delta(x-a)+\delta(x+a)]\,. \tag{5.4.61}$$

With these preliminaries out of the way we are ready for some examples.

5.5 Attractive Delta Function Potential

We want to find the bound states for the Hamiltonian

$$H = \frac{p^2}{2m} - \Lambda\delta(x) \quad , \quad \Lambda > 0\,. \tag{5.5.62}$$

The corresponding Schrödinger equation for the bound states is

$$H\psi = E\psi \tag{5.5.63}$$

with $E < 0$. Written out explicitly this reads:

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} - \Lambda\delta(x)\right]\psi(x) = E\psi(x) \tag{5.5.64}$$

or

$$\frac{d^2\psi}{dx^2} - \kappa^2\psi = -\alpha^2\delta(x)\psi(x) \tag{5.5.65}$$

where

$$\kappa^2 = -\frac{2mE}{\hbar^2} > 0 \tag{5.5.66}$$

and

$$\alpha^2 = \frac{2m\Lambda}{\hbar^2} > 0\,. \tag{5.5.67}$$

Thus, for $x \neq 0$ we have the Schrödinger equation

$$\frac{d^2\psi}{dx^2} - \kappa^2\psi = 0 \tag{5.5.68}$$

with the solutions

$$\begin{aligned} \psi(x) &= Ae^{-\kappa x} \quad x > 0 \\ \psi(x) &= Be^{\kappa x} \quad x < 0 \end{aligned} \tag{5.5.69}$$

where we have already used the fact that ψ must be square integrable and therefore vanish at $x \to \pm\infty$. We must also choose $\psi(x)$ to be continuous at $x = 0$. This gives

$$A = B \ . \tag{5.5.70}$$

However, the derivative of ψ, say ψ', is discontinuous at $x = 0$. This discontinuity can be computed directly from the Schrödinger equation (5.5.65). For this purpose we consider the expression

$$\begin{aligned} \lim_{\epsilon \to 0+} \int_{-\epsilon}^{\epsilon} \frac{d^2\psi}{dx^2}\,dx &= \psi'(0+) - \psi'(0-) \\ &= \lim_{\epsilon \to 0+} \left[\int_{-\epsilon}^{\epsilon} \kappa^2 \psi(x)\,dx - \alpha^2 \int_{-\epsilon}^{\epsilon} \psi(x)\delta(x)\,dx \right] \\ &= -\alpha^2 \psi(0) \end{aligned} \tag{5.5.71}$$

where we have used the fact that $\psi(0+) = \psi(0-)$. Thus,

$$\psi'(0+) - \psi'(0-) = -\alpha^2 \psi(0) \ . \tag{5.5.72}$$

Substituting the solutions equation (5.5.69) we get:

$$-\kappa A - \kappa A = -\alpha^2 A \tag{5.5.73}$$

or

$$\kappa = \alpha^2/2 = \frac{m\Lambda}{\hbar^2} \ . \tag{5.5.74}$$

But, since $E = -(\hbar^2\kappa^2)/2m$, we therefore get

$$E = -\frac{m\Lambda^2}{2\hbar^4} \ . \tag{5.5.75}$$

So the attractive delta function has one bound state with the above energy. The corresponding normalized wavefunction is

$$\psi = \sqrt{\kappa}e^{-\kappa|x|} \ . \tag{5.5.76}$$

The main point of this problem was to show how to handle the delta-function. We now consider the corresponding scattering problem, but for variety we choose a repulsive delta function.

5.6 Repulsive Delta Function Potential

In this case the potential is

$$V(x) = \Lambda\delta(x) \quad , \quad \Lambda > 0 \ . \tag{5.6.77}$$

As before the wavefunction ψ is continuous at $x = 0$ and has a discontinuity in its first derivative at $x = 0$. The Schrödinger equation can be rewritten to read

$$\frac{d^2\psi}{dx^2} + k^2\psi = \alpha^2\delta(x)\psi(x) \tag{5.6.78}$$

where $k^2 = 2mE/\hbar^2$ and $\alpha^2 = 2m\Lambda/\hbar^2$. If we call the solution for $x < 0$, $\psi_<$ and for $x > 0$, $\psi_>$ then the matching conditions at $x = 0$ are

$$\begin{aligned} \psi_<(0) &= \psi_>(0) \\ \psi'_<(0) &= \psi'_>(0) - \alpha^2\psi_>(0) \ . \end{aligned} \tag{5.6.79}$$

Since for $x \neq 0$ the Schrödinger equation reduces to the equation for a free particle, the solutions are $e^{\pm ikx}$. If we start with a particle incident from the left we have

$$\psi_<(x) = e^{ikx} + Re^{-ikx} \tag{5.6.80}$$

$$\psi_>(x) = Te^{ikx} \tag{5.6.81}$$

where we have used our freedom in choosing the amplitude of the incoming wave as well as the physical requirements that for $x > 0$ we have only a transmitted wave. Using the matching condition at $x = 0$ we now get

$$1 + R = T \tag{5.6.82}$$

$$ik - ikR = (ik - \alpha^2)T \ . \tag{5.6.83}$$

The solutions of these equations are

$$R = \frac{\alpha^2}{-\alpha^2 + 2ik} \tag{5.6.84}$$

$$T = \frac{2ik}{-\alpha^2 + 2ik} \ . \tag{5.6.85}$$

The transmission and reflection coefficients $\mathcal{T}$, $\mathcal{R}$ are then given by

$$\mathcal{T} = |T|^2 = \frac{4k^2}{4k^2 + \alpha^4} \tag{5.6.86}$$

$$\mathcal{R} = |R|^2 = \frac{\alpha^4}{4k^2 + \alpha^4} \ . \tag{5.6.87}$$

5.7 Square Well: Scattering and Phase Shifts

In section 4.2 we considered the finite square well

$$U(x) = \begin{cases} 0 & \text{if } |x| > a \\ -U & \text{if } |x| < a \end{cases} \tag{5.7.88}$$

and found all possible bound states. At the same time we saw that the invariance of this potential $V(x)$ under parity, that is

$$PV(x)P = V(-x) = V(x) \tag{5.7.89}$$

could be used to good advantage since the total Hamiltonian was then also invariant under parity. This allowed us to consider separately the positive parity (even) eigenfunctions and negative parity (odd) eigenfunctions. Such a separation into positive and negative parity states is also possible for scattering solutions and frequently simplifies the computations. However these solutions of definite parity do not correspond directly to the scattering solutions with the physical boundary conditions that we have been considering so far. Corresponding to the potential above, physical boundary conditions are:
For $x < -a$ we have an incoming wave (normalization 1) plus a reflected wave, or

$$\psi(x) = e^{ikx} + Re^{-ikx} \quad \text{for } x < -a\ . \tag{5.7.90}$$

For $x > a$ we have only a transmitted wave, or

$$\psi(x) = Te^{ikx} \qquad \text{for } x > a\ . \tag{5.7.91}$$

To make contact between these solutions and those of definite parity is fortunately quite easy.

As in section 4.2 we now define

$$k^2 = \frac{2mE}{\hbar^2}\ ,\quad K^2 = \frac{2m(E+U)}{\hbar^2}\ . \tag{5.7.92}$$

The Schrödinger equation then reads:

$$\begin{aligned} \frac{d^2\psi}{dx^2} &+ K^2\psi = 0 \quad |x| < a \\ \frac{d^2\psi}{dx^2} &+ k^2\psi = 0 \quad |x| > a\ . \end{aligned} \tag{5.7.93}$$

Choosing the normalization in the regions $|x| > a$ to be unity, we can write the solutions of definite parity as

$$\psi_+(x) = \begin{cases} \cos(kx - \delta_+) & \text{if } x < -a \\ A\cos Kx & \text{if } |x| < a \\ \cos(kx + \delta_+) & \text{if } x > a \end{cases} \tag{5.7.94}$$

and

$$\psi_-(x) = \begin{cases} i\sin(kx - \delta_-) & \text{if } x < -a \\ B\sin Kx & \text{if } |x| < a \\ i\sin(kx + \delta_-) & \text{if } x > a \end{cases} \tag{5.7.95}$$

where clearly we have

$$P\psi_\pm(x) = \psi_\pm(-x) = \pm\psi_\pm(x) \tag{5.7.96}$$

so that ψ_+ and ψ_- are respectively positive and negative parity solutions. The parameters $\delta_\pm$ are known as the phase shifts and reflect the presence of the potential. They contain all the information that the reflection and transmission amplitudes R and T contain. To see this we proceed as follows: For $x < -a$ we consider

$$e^{i\delta_+}\psi_+(x) + e^{i\delta_-}\psi_-(x) = e^{ikx} + \frac{1}{2}\left[e^{2i\delta_+} - e^{2i\delta_-}\right]e^{-ikx} . \tag{5.7.97}$$

Comparing this with the solution given by (5.7.90) for $x < -a$ we see that

$$R = \frac{1}{2}\left[e^{2i\delta_+} - e^{2i\delta_-}\right] . \tag{5.7.98}$$

Similarly for $x > a$ we consider

$$e^{i\delta_+}\psi_+(x) - e^{i\delta_-}\psi_-(x) = \frac{1}{2}\left[e^{2i\delta_+} + e^{2i\delta_-}\right]e^{ikx} \tag{5.7.99}$$

and comparing this with the solution given by (5.7.91) for $x > a$ we find

$$T = \frac{1}{2}\left[e^{2i\delta_+} + e^{2i\delta_-}\right] . \tag{5.7.100}$$

Thus, the problem is reduced to finding the phase shifts $\delta_\pm$.

To complete the solution and determine the phase shifts we have to match the wavefunctions and their slopes at $x = a$. There is a simplification that can be used due to the fact that we are not interested in the constants A, B in equations (5.7.94) and (5.7.95). Thus, instead of matching ψ and $d\psi/dx$ separately, we match the so-called logarithmic derivative $(1/\psi)(d\psi/dx)$. This has the effect of cancelling the unwanted constants.

From (5.7.94) we therefore get:

$$k\tan(ka + \delta_+) = K\tan Ka \tag{5.7.101}$$

and from (5.7.95) we get

$$k\cot(ka + \delta_-) = K\cot Ka \ . \tag{5.7.102}$$

Thus,

$$\delta_+ = \tan^{-1}\left[\frac{K}{k}\tan KA\right] - ka \tag{5.7.103}$$

$$\delta_- = \cot^{-1}\left[\frac{K}{k}\cot KA\right] - ka \ . \tag{5.7.104}$$

Inserting these results in equations (5.7.98) and (5.7.100) we obtain after some lengthy and tedious, simplifying algebra that

$$T = \frac{(k^2 - K^2)e^{-2ika}}{(k^2 + K^2) + ikK(\cot Ka - \tan Ka)} \tag{5.7.105}$$

and

$$R=\frac{ikK(\cot Ka+\tan Ka)e^{-2ika}}{(k^2+K^2)+ikK(\cot Ka-\tan Ka)} \ . \tag{5.7.106}$$

So as always $|T|^2+|R|^2=1$. This result is, in fact, obvious from equations (5.7.98) and (5.7.100) if $\delta_\pm$ are real. Also, as $U\to 0$ or, what amounts to the same thing, $K\to k$ we clearly see that $\delta_\pm\to 0$ and hence $R\to 0$ and $T\to 1$ as required by the physical situation. This is also obvious from equations (5.7.105) and (5.7.106). The expressions (5.7.105) and (5.7.106) for T and R could have been obtained directly by solving the Schrödinger equation (5.7.93) as we did in the previous examples. From the way we introduced the phase shifts, it would appear that they are only a computational tool. This is probably true for one-dimensional scattering problems. However in three-dimensional scattering problems (as we shall later see) the phase shifts have intrinsic physical interest, corresponding to quantities directly measurable in the laboratory. Thus, in anticipation of their future utility we have introduced them now.

5.8 Periodic Potentials

There are many physical situations which display discrete translational symmetry. Such situations are modelled by periodic potentials

$$V(x+a)=V(x) \ . \tag{5.8.107}$$

Iterating this equation we see that for $n=0,\pm1,\pm2,\cdots,N$

$$V(x+na)=V(x+(n-1)a)=\cdots=V(x) \ . \tag{5.8.108}$$

In practice, the number N is very large, of the order of Avagadro's number.

As a first step we show that if $u(x)$ is a solution of the corresponding Schrödinger equation then so is $u(x+a)$ and hence $u(x+na)$ $n=0,\pm1,\pm2,\cdots,N$. The Schrödinger equation for the solution $u(x)$ reads

$$-\frac{\hbar^2}{2m}\frac{d^2u(x)}{dx^2}+V(x)u(x)=Eu(x) \ . \tag{5.8.109}$$

Now consider the function $u(x+a)$. Then,

$$\begin{aligned}
&-\ \frac{\hbar^2}{2m}\frac{d^2u(x+a)}{dx^2}+V(x)u(x+a)\\
&=\ -\frac{\hbar^2}{2m}\frac{d^2u(x+a)}{d(x+a)^2}+V(x+a)u(x+a)\\
&=\ Eu(x+a) \ .
\end{aligned} \tag{5.8.110}$$

This shows that that $u(x+a)$ is again a solution. Now if two linearly independent solutions of (5.8.109) are $u_1(x)$ and $u_2(x)$ then we must have that the two linearly independent solutions $u_1(x+a)$ and $u_2(x+a)$ are linear combinations of $u_1(x)$ and $u_2(x)$. This means that we have

$$\begin{aligned}
u_1(x+a) &= c_{11}u_1(x)+c_{12}u_2(x)\\
u_2(x+a) &= c_{21}u_1(x)+c_{22}u_2(x) \ .
\end{aligned} \tag{5.8.111}$$

5.8.1 Floquet's Theorem

We now show that the coefficients c_{ij} may be chosen such that there are solutions v_i , $i = 1, 2$ such that

$$\begin{aligned} v_1(x+a) &= \lambda_1 v_1(x) \\ v_2(x+a) &= \lambda_2 v_2(x) \end{aligned} \tag{5.8.112}$$

where λ_i are constants such that

$$\lambda_1 \lambda_2 = 1 \ . \tag{5.8.113}$$

It furthermore follows immediately that

$$v_i(x+na) = \lambda_i^n\, v_i(x) \quad n = 0\,, \pm 1\,, \pm 2\,, \cdots N \ . \tag{5.8.114}$$

This is Floquet's theorem. The proof is straightforward. We simply use the fact that the functions $v_i(x)$ are again linear combinations of $u_i(x)$ and write out (5.8.111). The result is

$$\begin{aligned} v_i(x+a) &= A_i u_1(x+a) + B_i u_2(x+a) \\ &= A_i\left[c_{11}u_1(x) + c_{12}u_2(x)\right] + B_i\left[c_{21}u_1(x) + c_{22}u_2(x)\right] \\ &= \lambda_i\, v_i(x) \\ &= \lambda_i\left[A_i u_1(x) + B_i u_2(x)\right] \ . \end{aligned} \tag{5.8.115}$$

For this equation to hold requires that

$$\begin{aligned} A_i c_{11} + B_i c_{21} &= \lambda_i\, A_i \\ A_i c_{12} + B_i c_{22} &= \lambda_i\, B_i \ . \end{aligned} \tag{5.8.116}$$

For a non-trivial solution for A_i , B_i we need that the determinant of the coefficients vanish.

$$\begin{vmatrix} c_{11} - \lambda_i & c_{21} \\ c_{12} & c_{22} - \lambda_i \end{vmatrix} = 0 \ . \tag{5.8.117}$$

This quadratic equation clearly yields two solutions λ_1 and λ_2. Thus, we also have two solutions v_1 and v_2.

The Wronskian W (see problem 4.11) of these two solutions

$$W(x) = v_1 \frac{dv_2}{dx} - v_2 \frac{dv_1}{dx} \tag{5.8.118}$$

is, as shown in problem 4.11, a constant and using (5.8.112) clearly also satisfies the relationship

$$W(x+a) = \lambda_1 \lambda_2\, W(x) \ . \tag{5.8.119}$$

This means that

$$\lambda_1 \lambda_2 = 1 \tag{5.8.120}$$

as was claimed at the beginning of this section.

5.8.2 Bloch's Theorem

Bloch gave a more detailed description of the solutions we have just discussed. His results are of far-reaching importance for studies of crystalline structures. We now describe his findings.

Suppose $\lambda_1 > 1$ and therefore $\lambda_2 < 1$. In this case we find by repeated use of (5.8.114) that

$$v_1(x+na) = \lambda_1^n \, v_1(x) \quad n = 0\,, \pm 1\,, \pm 2\,, \cdots \tag{5.8.121}$$

and for $n \to +\infty$ the amplitude of $v_1(x+na)$ blows up. This means that for $x \to +\infty$ $v_1(x)$ blows up and hence cannot be normalized to 1. Similarly, applying this result to $v_2(x)$ we find that

$$v_2(x-na) = \lambda_2^{-n} \, v_2(x) \quad n = 0\,, 1\,, 2\,, \cdots, N \tag{5.8.122}$$

and for $n \to +\infty$ the amplitude of $v_2(x-na)$ blows up. So again this solution can not be normalized. This shows that for an acceptable physical solution we may only have

$$|\lambda_i| = 1 \ . \tag{5.8.123}$$

In this case we can choose

$$\lambda_1 = e^{iKa} \quad \lambda_2 = e^{-iKa} \tag{5.8.124}$$

with K real and such that $-\pi a \leq K \leq \pi a$ since $\exp(i2\pi n) = 1$. This means that the physically acceptable solutions are such that

$$v(x+na) = e^{inKa} \, v(x) \ . \tag{5.8.125}$$

This equation is only possible if the $v(x)$ are of the form

$$v(x) = e^{iKx} \, v_K(x) \tag{5.8.126}$$

with

$$v_K(x+a) = v_K(x) \ . \tag{5.8.127}$$

This is the celebrated *Bloch's Theorem.* The value of K follows from the boundary conditions that we apply to the extreme ends of the crystal. In practice, we usually apply periodic boundary conditions to the end of the lattic and the beginning. In other words, we think of the one-dimensional lattice as a circle. Thus, we require that

$$v(x+Na) = v(x) \ . \tag{5.8.128}$$

Using (5.8.125) this becomes

$$e^{iNKa} \, v(x) = v(x) \ . \tag{5.8.129}$$

Thus,

$$e^{iNKa} = 1 \tag{5.8.130}$$

and

$$K = \frac{2\pi n}{Na} \quad n = 0\,,\pm 1\,,\pm 2\,,\cdots \tag{5.8.131}$$

The utility of Bloch's theorem is that we need solve the Schrödinger equation in only one of the intervals say $0 < x < a$ and then use the appropropriate boundary conditions.

5.9 The Kronig-Penney Problem

To illustrate the use of the results discussed above we now solve the following simple model known as the *Kronig-Penney Problem*. For the potential we choose

$$V(x) = V_0 a \sum_{n=-\infty}^{n=\infty} \delta(x + na) \ . \tag{5.9.132}$$

This potential is also known as a *Dirac comb*. After a change of variables so that we can introduce the wavenumber k and the dimensionless variable U

$$E = \frac{\hbar^2 k^2}{2m} \geq 0 \quad , \quad U = \frac{2mV_0 a^2}{\hbar^2} \tag{5.9.133}$$

the Schrödinger equation becomes

$$\frac{d^2\psi}{dx^2} - \frac{U}{a} \sum_{n=-\infty}^{n=\infty} \delta(x + na)\,\psi + k^2\psi = 0\ . \tag{5.9.134}$$

We next apply Bloch's theorem and look for solutions valid in the interval $0 < x < a$. These are clearly solutions of

$$\frac{d^2 u}{dx^2} + k^2 u = 0 \tag{5.9.135}$$

and are of the form

$$u(x) = A \sin kx + B \cos kx \quad 0 < x < a \ . \tag{5.9.136}$$

It now follows from Bloch's theorem that for $a < x < 2a$ the solution is given by

$$u(x) = e^{iKa}\left[A \sin k(x-a) + B \cos k(x-a)\right] \quad a < x < 2a \ . \tag{5.9.137}$$

The boundary conditions at $x = a$ are
1) $u(x)$ is continuous so that

$$u(a+0) = u(a-0) \ . \tag{5.9.138}$$

2) The second boundary condition follows from integrating equation (5.9.134) about $x = a$ and yields a discontinuity in the first derivative of $u(x)$. The result is

$$u'(a+0) - u'(a-0) = \frac{U}{a}\,u(a) \ . \tag{5.9.139}$$

Applying these conditions we get

$$\begin{aligned} B\,e^{iKa} &= A\,\sin ka + B\,\cos ka \\ kA\,e^{iKa} - k[A\,\cos ka - B\,\sin ka] &= \frac{U}{a} B\,e^{iKa} \ . \end{aligned} \tag{5.9.140}$$

For a nontrivial solution for A and B requires that the determinant of their coefficients vanish. Thus, we need

$$\det \begin{vmatrix} \sin ka & \cos ka - e^{iKa} \\ e^{iKa} - \cos ka & \sin ka - \frac{U}{ka}\,e^{iKa} \end{vmatrix} = 0 \ . \tag{5.9.141}$$

The result is

$$\sin^2 ka - \frac{U}{ka}\,e^{iKa} \sin ka + \left(\cos ka - e^{iKa}\right)^2 = 0 \ . \tag{5.9.142}$$

After a little algebra this reduces to

$$\cos Ka = \cos ka + \frac{U}{2ka} \sin ka \ . \tag{5.9.143}$$

The energy eigenvalues are given by (5.9.133)

$$E = \frac{\hbar^2 k^2}{2m} \tag{5.9.144}$$

and the allowed values of k are to be determined from (5.9.143). In this case we see that we get *bands* of energy. In fact all values of k such that

$$|\cos ka + \frac{U}{2ka} \sin ka| \le 1 \tag{5.9.145}$$

are allowed. Thus, we have regions of allowed k values (energy bands) with *gaps*, given by

$$|\cos ka + \frac{U}{2ka} \sin ka| \ge 1 \tag{5.9.146}$$

between them. These gaps represent regions of forbidden energies. Due to the very large value of N almost any energy is permitted within the band. It is the existence of the energy gaps that accounts for the very large difference in conductivity between insulators, semi-conductors and conductors. This model serves to illustrate the existence of such bands.

So far we have been solving very concretely specified problems. The purpose of this has been two-fold: to develop our technical abilities and to build up our intuition. We now leave the discussion of specific problems for a while to embark on a consideration of the formal aspects of quantum mechanics. In this process we shall develop a certain mathematical language which may, at first, seem artificial but is, in fact, very important if not essential for a clear formulation of quantum mechanics.

5.10 Problems

5.1 A wavefunction

$$\psi(t,x) = \int_{-\infty}^{\infty} f(k) e^{i[kx-\omega(k)t]}\, dk$$

is normalized such that

$$\int_{-\infty}^{\infty} |\psi(t,x)|^2\, dx = 1 .$$

Assume $f(k)$ is a smooth function vanishing rapidly at infinity. Show that the velocity of the centre of mass $\bar{x}$ defined by

$$\bar{x} = \int_{-\infty}^{\infty} x|\psi(t,x)|^2\, dx$$

is given by

$$\frac{d\bar{x}}{dt} = 2\pi \int_{-\infty}^{\infty} \frac{d\omega}{dk} |f(k)|^2\, dk .$$

Hint: Use the fact that

$$\int_{-\infty}^{\infty} e^{i(k-q)x}\, dx = 2\pi\delta(k-q) .$$

Show furthermore that if $\omega(k)$ is adequately approximated by

$$\omega(k) \approx \omega(0) + (k-k_0) \left.\frac{d\omega}{dk}\right|_{k_0}$$

then

$$\frac{d\bar{x}}{dt} \approx \left.\frac{d\omega}{dk}\right|_{k_0} ,$$

(the group velocity).

5.2 Compute the probability density ρ and the current j for the wavefunction in the previous problem 5.1. What are the possible forms of $\omega(k)$ such that the equation of continuity is satisfied? Explain this result recalling that $E = \hbar\omega$ and $p = \hbar k$.

5.3 A potential sometimes used in molecular physics is the so-called 6-12 potential

$$V(x) = \epsilon_0 \left[\left(\frac{\sigma}{x}\right)^{12} - 2\left(\frac{\sigma}{x}\right)^{6} \right] \quad 0 \leq x < \infty .$$

a) Sketch this potential and indicate the values of the most interesting points in the diagram.
b) Estimate the ground state (lowest) energy eigenvalue for a particle of mass m bound in this potential.

5.4 A particle is in the potential

$$V(x) = V_0 \exp[ax^2/2] \ .$$

Estimate the energy of the 2 lowest eigenstates. What are the parities of these states?

5.5 Find the wave-function for a particle at rest at the origin of a coordinate system fixed in space.

5.6 Calculate the transmission probability for a particle of mass m incident on a potential

$$V(x) = \Lambda[\delta(x+a) + \delta(x-a)] \ .$$

Compute also the phase shifts.

5.7 Complete all the steps in going from (5.7.98), (5.7.100), (5.7.103), (5.7.104) to (5.7.105), and (5.7.106).

5.8 Using the representation

$$\delta(x) = \lim_{\epsilon\to 0} h_\epsilon(x)$$

where

$$h_\epsilon(x) = \begin{cases} \frac{1}{2\epsilon} & \text{if } -\epsilon < x < \epsilon \\ 0 & \text{if } -\epsilon > x > \epsilon \end{cases} .$$

Verify equations (5.4.57) to (5.4.61).
Hint: Integrate both sides of the equation with a well-behaved function $f(x)$.

5.9 Repeat problem 5.8 using the representation (5.4.52).

5.10 Find the energy levels and normalized wavefunctions of the stationary states of a particle moving in the potential

$$V(x) = \begin{cases} \infty & \text{if } x < 0 \\ \frac{1}{2}kx^2 & \text{if } x > 0 \end{cases} .$$

Compare the zero point energy in this potential with that for the simple harmonic oscillator with the same force constant k. Explain any differences you find.

Bibliography

[5.1] All the references for Chapter 4 apply. In addition, it may be useful to consult:
S. Flügge, *Practical Quantum Mechanics* - Springer-Verlag, New York, Heidelberg, Berlin (1974).
This book is a collection of very many worked examples together with a discussion of the physics of these examples.

[5.2] A. Sommerfeld, *Wave Mechanics* - Dutton, New York (1929).

[5.3] R. de Kronig and L.W. Penney, Proc. Roy. Soc.,London, **130**, 499,(1931).

Chapter 6

Mathematical Foundations

6.1 Introduction

In this chapter we present a mathematical interlude to provide a more formal language which will allow us to set the mathematical foundations of quantum mechanics. Sections 6.2 and 6.3 contain a discussion of Hilbert space and should be read by all those not familiar with this concept. The remaining sections are somewhat more mathematical in nature and are provided for those who desire more rigour.

In section 6.4 we discuss linear operators in Hilbert space and introduce the concept of self-adjointness. The Cayley transform is introduced in section 6.5 and used to classify all self-adjoint extensions of a symmetric operator. Section 6.6 is devoted to some examples illustrating the results of section 6.5. More examples are also provided in section 6.7.

6.2 Geometry of Hilbert Space

The language we have used so far is one of wave-functions or state vectors and operators on these wave-functions. There is a ready-made mathematical language for this. This is the language of Hilbert space and quantum mechanics is naturally formulated in Hilbert space. Actually, for practical purposes Hilbert space is too small and the appropriate generalization is to a so-called rigged Hilbert space. We shall ignore this for the time being and discuss it briefly in Chapter 8, where we shall also give references for those who are interested in more of the details. What is Hilbert space? First of all it is a vector space analogous to the usual Euclidean spaces $\mathcal{E}_3$ or $\mathcal{E}_n$, however unlike $\mathcal{E}_3$ which is 3-dimensional and $\mathcal{E}_n$ which is n-dimensional, Hilbert space is ∞-dimensional. As a conceptual model of Hilbert space $\mathcal{H}$ one can think of taking some representation of a vector in $\mathcal{E}_n$, say $(a_1, \ldots, a_n)$ and writing it $(a_1, \ldots, a_n, 0, 0 \ldots, 0)$. Then, letting n increase without limit we arrive at the notion of $\mathcal{H}$. However, unlike all ordered n-tuples which automatically can be considered in $\mathcal{E}_n$, not all infinite ordered sequences can be considered to belong to $\mathcal{H}$. The reason for this is that not all of them have finite "length" and we wish to include in $\mathcal{H}$ only elements of finite length. This is tantamount to saying in the language of wavefunctions that we want the wavefunction to be square integrable. We start by listing those properties of $\mathcal{E}_n$ which remain true in the transition to $\mathcal{H}$. The elements of the space are vectors f on which certain operations are defined.

1) Scalar multiplication : If $f \in \mathcal{H}$ and λ is a complex number then $\lambda f \in \mathcal{H}$.

2) Addition: If f_1 and $f_2 \in \mathcal{H}$ then $f_1 + f_2 \in \mathcal{H}$. Thus, combining 1) and 2) we see that all finite linear combinations of elements in $\mathcal{H}$ belong to $\mathcal{H}$.

3) Inner Product: In $\mathcal{H}$ there is defined an inner product (f, g) for all $f, g \in \mathcal{H}$. This inner product maps elements of $\mathcal{H}$ into complex numbers and satisfies the following conditions.
a)

$$(f,g) = (g,f)^* \tag{6.2.1}$$

where the star means "complex conjugation".
b)

$$(\lambda_1 f, \lambda_2 g) = \lambda_1^* \lambda_2 (f,g) \tag{6.2.2}$$

where λ_1 and λ_2 are complex numbers. We sometimes also use a bar over a number to indicate complex conjugation.
c)

$$\begin{aligned} (f_1 + f_2, g) &= (f_1, g) + (f_2, g) \\ (f, g_1 + g_2) &= (f, g_1) + (f, g_2) \ . \end{aligned} \tag{6.2.3}$$

d)

$$|(f,g)|^2 \leq (f,f)(g,g) \ . \tag{6.2.4}$$

This last inequality is the Schwarz inequality. We shall later derive this.

In terms of the inner product we define the "length" or norm of a vector in the usual way by

$$\| f \|^2 = (f,f) \ . \tag{6.2.5}$$

It is also possible to define orthogonality using the inner product. Thus, f is orthogonal to g if and only if (abbreviated iff)

$$(f,g) = 0 \ . \tag{6.2.6}$$

A set of vectors $\{f_i\}$ is orthonormal that is orthogonal and normal iff

$$(f_i, f_j) = \delta_{ij} \ . \tag{6.2.7}$$

We now come to some of the differences. For this we need two definitions.

Definition: A set $\{f_i\}$ of vectors is *complete* iff any vector in $\mathcal{H}$ can be written as a linear combination of vectors from the set $\{f_i\}$. A complete set of orthonormal vectors forms a *basis*.

Example:
Consider the Euclidean space $\mathcal{E}_3$ and choose three orthonormal vectors $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$. Thus, $(\hat{e}_i, \hat{e}_j) = \delta_{ij}$. Then any vector $f \in \mathcal{E}_3$ can be written

$$f = \sum_{i=1}^{3} \lambda_i \hat{e}_i \tag{6.2.8}$$

and in fact

$$\lambda_i = (\hat{e}_i, f) \, . \tag{6.2.9}$$

The $\hat{e}_i$ are obviously complete and form a basis in $\mathcal{E}_3$. On the other hand, if we choose just two of these vectors, say $\hat{e}_1$ and $\hat{e}_2$, they do not form a complete set since for example any vector with a component along $\hat{e}_3$ cannot be expressed in terms of just the first two.

The λ_i are usually called the components of the vector. If we then agree to keep the basis fixed, we can suppress the basis vectors and write $f = (\lambda_1, \lambda_2, \lambda_2)$. In this manner we establish a one-one correspondence between vectors in $\mathcal{E}_3$ and ordered triplets. The norm of f is given by

$$\begin{aligned} \| f \|^2 &= \sum_{i,j=1}^{3} \lambda_i^* \lambda_j (\hat{e}_i, \hat{e}_j) \\ &= \sum_{i,j=1}^{3} \lambda_i^* \lambda_j \delta_{ij} \\ &= \sum_{i=1}^{3} |\lambda_i|^2 \, . \end{aligned} \tag{6.2.10}$$

This formalism above is all exceedingly trivial and you may wonder why bother. The reason is to establish a precise formalism so that when the situation becomes complicated we can rely on the formalism and not just our intuition.

Another use for the word *complete* is in the description of a vector space. This concept again is trivial for $\mathcal{E}_n$ but is non-trivial for $\mathcal{H}$. Consider a sequence of vectors $f_1, f_2, f_3, \ldots$. Furthermore, suppose that for every $\epsilon > 0$ we can find an n such that for any finite m

$$\| f_{n+m} - f_n \| < \epsilon \, . \tag{6.2.11}$$

This is just a statement of the Cauchy criterion for convergence using the norm $\| \cdot \|$ rather than the absolute value as is the usual case for numerical sequences. We call such a sequence a *Cauchy sequence*. Now if $\mathcal{H}$ in this case is finite dimensional say $\mathcal{E}_n$ then it is trivial to show that the limit of the sequence exists and is a vector in $\mathcal{E}_n$. This property that all Cauchy sequences have a limit in $\mathcal{E}_n$ is stated by saying that $\mathcal{E}_n$ is complete. In fact all finite dimensional vector spaces are complete. This is also true for ∞-dimensional Hilbert spaces. There is a deep theorem of analysis known as the *Riesz-Fischer Theorem* which states that the space used by us (called $\mathcal{L}_2$ by mathematicians) is complete and thus a Hilbert space.

6.3 $\mathcal{L}_2$: A Model Hilbert Space

The elements of $\mathcal{L}_2$ are square-integrable complex-valued functions $f(x)$ of a real variable x. More generally x is a vector in some real finite vector space so that f is a function of n real variables. This generalization has no effect whatever on the ensuing statements and so we ignore it. The norm $\|\cdot\|$ in $\mathcal{L}_2$ is defined by

$$\| f \|^2 = (f, f) = \int f^*(x) f(x)\, dx \,. \tag{6.3.12}$$

The range of integration in (6.3.12) is over the full range of the variable x. Thus, if x is unrestricted, the integral runs from $-\infty$ to ∞. It is trivial to check our first two conditions for elements in $\mathcal{L}_2$. Thus,

$$f \in \mathcal{L}_2 \;\; \Rightarrow \;\; \lambda f \in \mathcal{L}_2 \tag{6.3.13}$$

and

$$f_1 \,,\, f_2 \in \mathcal{L}_2 \;\; \Rightarrow \;\; f_1 + f_2 \in \mathcal{L}_2 \;. \tag{6.3.14}$$

Furthermore, as defined by (6.3.12) the inner product obviously satisfies conditions a), b) and c).

a)

$$(f, g) = \int f^* g\, dx = \left(\int f g^*\, dx \right)^* = (g, f)^* \tag{6.3.15}$$

b)

$$(\lambda_1 f, \lambda_2 g) = \int (\lambda_1 f)^* (\lambda_2 g)\, dx = \lambda_1^* \lambda_2 \int f^* g\, dx = \lambda_1^* \lambda_2 (f, g) \tag{6.3.16}$$

c)

$$\begin{aligned} (f_1 + f_2, g) &= \int (f_1 + f_2)^* g\, dx = \int f_1^* g\, dx + \int f_2^* g\, dx \\ &= (f_1, g) + (f_2, g) \end{aligned} \tag{6.3.17}$$

and

$$\begin{aligned} (f, g_1 + g_2) &= \int f^* (g_1 + g_2)\, dx = \int f^* g_1\, dx + \int f^* g_2\, dx \\ &= (f, g_1) + (f, g_2) \;. \end{aligned} \tag{6.3.18}$$

The only condition left to verify on the inner product is the Schwarz inequality. To do this consider the vector

$$h = f + \lambda(g, f)g \tag{6.3.19}$$

where $f, g \in \mathcal{L}_2$ and λ is a real number. Then,

$$(h, h) = \| h \|^2 \geq 0 \ . \tag{6.3.20}$$

Therefore,

$$\begin{aligned} 0 &\leq (f + \lambda(g, f)g, f + \lambda(g, f)g) \\ &= (f, f) + 2\lambda|(f, g)|^2 + \lambda^2|(f, g)|^2(g, g) \ . \end{aligned} \tag{6.3.21}$$

This means that the quadratic polynomial in λ cannot have two real distinct zeros and hence that the discriminant is negative, giving

$$|(f, g)|^4 - |(f, g)|^2(f, f)(g, g) \leq 0 \ . \tag{6.3.22}$$

The equality sign obviously applies when $(f, g) = 0$ or $f = \mu g$. Thus, even if $(f, g) = 0$, we get

$$|(f, g)|^2 \leq (f, f)(g, g) \tag{6.3.23}$$

as required.

Orthogonality is still given by: f is orthogonal to g iff

$$(f, g) = 0 \ . \tag{6.3.24}$$

An example of two orthogonal vectors in $\mathcal{L}_2$ is:

$$f = \frac{1}{\pi^{1/4}}\, e^{-x^2/2} \quad g = \frac{2x}{\pi^{1/4}}\, e^{-x^2/2} \tag{6.3.25}$$

where the range of intergration is $(-\infty\, , \infty)$. The functions displayed are the first two hermite functions. As we see later, it is a general fact that eigenfunctions corresponding to different eigenvalues are orthogonal. In fact for physical Hamiltonians, the eigenfunctions properly normalized can be taken as a basis. This is an important fact since it implies that these eigenvectors form a complete set. Although we do not prove the Riesz-Fischer Theorem we restate it here. Consider a sequence of functions $f_1, f_2, \ldots$ all of which belong to $\mathcal{L}_2$. Furthermore, let this be a Cauchy sequence. This means that given any $\epsilon > 0$ we can find an $n > 0$ such that

$$\left[\int |f_{n+m} - f_n|^2\, dx\right]^{1/2} < \epsilon \ . \tag{6.3.26}$$

Then the Riesz-Fischer Theorem asserts that
a) $\lim_{m\to\infty} f_m(x) = f(x)$
exists and, most important,

b) $f(x) \in \mathcal{L}_2$.
That is,

$$\int |f(x)|^2 \, dx < \infty \ . \tag{6.3.27}$$

This guarantees that we can take limits of sequences in $\mathcal{L}_2$. Clearly $\mathcal{L}_2$ is the model for the quantum mechanical Hilbert spaces. All Hilbert spaces are complete by definition and so all limits of Cauchy sequences are again elements of the space. We shall see that the inner product plays an exceedingly important role in the physical interpretation of quantum mechanics. There is one more technical point, namely that for any set of linearly independent vectors $\{f_j\}$ it is possible to construct an orthonormal set of vectors $\{e_j\}$ which span the same space. The orthogonalization process is called the *Schmidt Orthogonalization Procedure.* The proof is by construction. Choose one of the f_j say f_1. Then,

$$e_1 = \frac{f_1}{\| f_1 \|} \ . \tag{6.3.28}$$

Now form

$$g_2 = f_2 - (e_1, f_2)e_1 \tag{6.3.29}$$

and

$$e_2 = \frac{g_2}{\| g_2 \|} \ . \tag{6.3.30}$$

Clearly, e_1 and e_2 are orthonormal. Next form

$$g_3 = f_3 - (e_1, f_3)e_1 - (e_2, f_3)e_2 \tag{6.3.31}$$

and

$$e_3 = \frac{g_3}{\| g_3 \|} \ . \tag{6.3.32}$$

The process is now obvious.

We next turn to another aspect of Hilbert space, namely operators.

6.4 Operators on Hilbert Space: Mainly Definitions

An operator on Hilbert space is a mapping which maps certain elements of $\mathcal{H}$ into $\mathcal{H}$. Thus, if A is an operator with *domain* $D_A \subset \mathcal{H}$ then for all $f \in D_A$

$$g = Af \in \mathcal{H} \ . \tag{6.4.33}$$

The domain D_A consists simply of all those vectors in $\mathcal{H}$ such that the result of operating with A on a vector in D_A is again a vector in $\mathcal{H}$. Thus, for example, if $\mathcal{H}$ is the space $\mathcal{L}_2$ and A is the operator x^2 (multiplication by x^2) then

$$(Af)(x) = x^2 f(x) \ . \tag{6.4.34}$$

Clearly even if $f \in \mathcal{L}_2$ not all functions $x^2 f(x)$ are in $\mathcal{L}_2$. For example

$$f(x) = (x^2 + a^2)^{-m} \tag{6.4.35}$$

is in $\mathcal{L}_2$ for $\Re(m) > 1/4$. But, $x^2 f(x) = x^2(x^2 + a^2)^{-m}$ is not in $\mathcal{L}_2$ unless $\Re(m) > 5/4$.

Again we shall only be interested in linear operators. Thus,

$$A(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 A f_1 + \lambda_2 A f_2 \ . \tag{6.4.36}$$

For example, the operator x^2 defined above is linear and so are the operators

$$pf = \frac{\hbar}{i} \frac{df}{dx} \tag{6.4.37}$$

and

$$(Kf)(x) = \int_{-\infty}^{\infty} K(x, y) f(y)\, dy \ . \tag{6.4.38}$$

On the other hand, $\log f$ and $\sqrt{f}$ are definitely not linear operators acting on the function f . Linear operators are also familiar in finite dimensional vector spaces. They are usually represented by matrices in this case. There is an analogous representation for operators in Hilbert space. Formulated in this way quantum mechanics is called matrix mechanics to distinguish it from the Schrödinger formulation or wave mechanics. Both formulations are just two different mathematical ways of looking at the same thing. We examine matrix mechanics after we have developed all the necessary mathematical machinery. To illustrate the matrix operator formalism we first derive the form of the most general linear operator on a finite vector space say $\mathcal{E}_n$. Let A be such an operator. Call

$$g = Af \tag{6.4.39}$$

and consider taking for f different elements of a basis set $\{e_i\}$. Thus, let

$$g_i = A e_i \ . \tag{6.4.40}$$

Then writing

$$f = \sum_i \lambda_i e_i \tag{6.4.41}$$

and

$$g = \sum_i \mu_i e_i \tag{6.4.42}$$

we get by linearity:

$$\sum_i \mu_i e_i = \sum_i \lambda_i A e_i = \sum_i \lambda_i g_i \ . \tag{6.4.43}$$

Now using

$$(e_i, e_j) = \delta_{ij} \tag{6.4.44}$$

and taking inner products in (6.4.43) we get:

$$\mu_j = \sum_i \lambda_i (e_j, Ae_i) = \sum_i \lambda_i (e_j, g_i) \ . \tag{6.4.45}$$

This means that the operator A is completely determined in this basis by the matrix of numbers

$$A_{ji} = (e_j, Ae_i) \ . \tag{6.4.46}$$

Conversely if we are given a matrix $(n \times n)$ then it can always be used to define a linear operator according to (6.4.46). Thus, as stated previously, the most general linear operator on $\mathcal{E}_n$ can be considered to be an $(n \times n)$ matrix. With only some attention to details the same argument will go through for an ∞-dimensional Hilbert space. Just as a matrix algebra is possible, so an algebra of linear operators is generally possible. It is only necessary to pay due attention to such things as domains of the operators. Let A be an operator on $\mathcal{H}$ with domain D_A. Then λA is also an operator on $\mathcal{H}$ with domain D_A and acts as follows:

$$(\lambda A)f = \lambda(Af) \ . \tag{6.4.47}$$

This is almost too obvious. If A, B are operators with domains D_A and D_B respectively, then $A + B$ is an operator with domain $D_A \cap D_B$ defined by

$$(A + B)f = Af + Bf \ . \tag{6.4.48}$$

The *range* R_A of an operator A is defined as the set of all vectors obtained by operating with A on elements in D_A. Symbolically,

$$R_A = AD_A \ . \tag{6.4.49}$$

Then if A, B are operators with domains D_A and D_B and $R_A \subset D_B$ we can define the product operator BA according to

$$(BA)f = B(Af) \ . \tag{6.4.50}$$

This is well defined since by assumption

$$f \in D_A \ . \tag{6.4.51}$$

Therefore

$$Af \in R_A \subset D_B \tag{6.4.52}$$

and hence

$$Af \in D_B \ . \tag{6.4.53}$$

Conversely if $R_B \subset D_A$ we can define the product

$$(AB)f = A(Bf) \ . \tag{6.4.54}$$

This points out the interesting possibility that although BA may exist as an operator AB might not, and conversely. Another property which many operators of physical interest possess is hermiticity. Actually the interesting property is self-adjointness and we shall examine these two properties in some detail to bring out the difference. First we need some definitions. Let A be an operator on $\mathcal{H}$ and $f \in D_A$. Then consider the expression (g, Af). If for some $g \in \mathcal{H}$ we find that there is an $h \in \mathcal{H}$ such that

$$(g, Af) = (h, f) \tag{6.4.55}$$

for all $f \in D_A$ then we define the adjoint operator $A^\dagger$ of A by

$$h = A^\dagger g \tag{6.4.56}$$

with domain $D_{A^\dagger}$ = the set of all g for which (6.4.55) holds. In that case we can write

$$(g, Af) = (A^\dagger g, f) \ . \tag{6.4.57}$$

Note that the element h in (6.4.55) is defined uniquely by g if the domain D_A contains sufficiently many vectors. The precise statement of this is that D_A is *dense* in $\mathcal{H}$. For our purposes a set in $\mathcal{H}$ is dense if any element in $\mathcal{H}$ can be approximated arbitrarily closely by an element from this set. Thus, D_A is dense if for any $f \in \mathcal{H}$ there exists a $g \in D_A$ such that given

$$\epsilon > 0 \quad \| f - g \| < \epsilon \ . \tag{6.4.58}$$

In this case the proof that h is unique is trivial. For, assume there is another such vector h'. Then,

$$(g, Af) = (h', f) \tag{6.4.59}$$

as well. Combining this with (6.4.55) we get

$$(h - h', f) = 0 \ . \tag{6.4.60}$$

Thus, $h - h'$ is orthogonal to every vector in D_A. But D_A is dense in $\mathcal{H}$ so that for any vector $g \in \mathcal{H}$

$$|(h - h', g)| < \epsilon \ . \tag{6.4.61}$$

This is possible only if $h = h'$. The adjoint operator is also a linear operator as is immediately obvious. Now again let A be an operator in $\mathcal{H}$ with domain D_A then A is *hermitian* if for all $f, g \in D_A$

$$(Af, g) = (f, Ag) \ . \tag{6.4.62}$$

An operator A is *symmetric* if it is hermitian and its domain of definition D_A is dense in $\mathcal{H}$. From the definition of $A^\dagger$ it then follows that for a symmetric operator A

$$D_A \subset D_{A^\dagger} \tag{6.4.63}$$

as we show for an example. If in addition

$$D_A = D_{A^\dagger} \tag{6.4.64}$$

or as this implies

$$A = A^\dagger \ . \tag{6.4.65}$$

Then, A is *self-adjoint.* To make this less abstract consider the momentum operator

$$p = \frac{\hbar}{i}\frac{d}{dx} \tag{6.4.66}$$

defined on the Hilbert space $\mathcal{L}_2(a,b)$ of functions square-integrable on the interval $[a,b]$. As domain of this operator we choose

$$D_p = \left\{ f \in \mathcal{L}_2(a,b) \mid \frac{df}{dx} \text{ is bounded on } (a,b)\ ,\ f(a) = f(b) = 0 \right\} \ . \tag{6.4.67}$$

With this definition it is easy to see that p is hermitian and in fact symmetric. The domain D_p is dense in $\mathcal{L}_2(a,b)$. Thus, if p is hermitian, it is symmetric. To see hermiticity let

$$f \in D_p \quad , \quad g \in D_p \ . \tag{6.4.68}$$

Then,

$$\begin{aligned} (f,pg) &= \int_a^b f^*(x)\frac{\hbar}{i}\frac{dg(x)}{dx}\,dx \\ &= \frac{\hbar}{i}\,f^*(x)g(x)|_a^b + \int_a^b \left(\frac{\hbar}{i}\frac{df(x)}{dx}\right)^* g(x)\,dx \ . \end{aligned} \tag{6.4.69}$$

Since $f^*(a) = f^*(b) = g(a) = g(b) = 0$, the term obtained from integration by parts vanishes and so we have

$$(f,pg) = (pf,g) \ . \tag{6.4.70}$$

Thus, p is hermitian (symmetric) as claimed. On the other hand $p \neq p^\dagger$ since, as we now show, the domain $D_{p^\dagger}$ of $p^\dagger$ is much larger than the domain D_p of p, i.e. $D_p \subset D_{p^\dagger}$, but $D_p \neq D_{p^\dagger}$. This means that D_p is a proper subset of $D_{p^\dagger}$. To see this consider any $g \in D_p$ and let f be any function whose derivative is bounded over (a,b) and such that

$$f(b) = e^{i\theta} f(a) \tag{6.4.71}$$

where θ is a constant. Then by a computation, identical to the one above, we again find

$$(f, pg) = (pf, g) \ . \tag{6.4.72}$$

Thus, as a differential operator

$$p^\dagger = \frac{\hbar}{i}\frac{d}{dx} \tag{6.4.73}$$

but the domain of $p^\dagger$ is larger than the domain of p. It is furthermore easy to check that if we define

$$D_{p^\dagger} = \left\{ f \in \mathcal{L}_2(a,b) \ \middle| \ \frac{df}{dx} \text{ is bounded on } (a,b) \text{ and } f(b) = e^{i\theta} f(a) \right\} \tag{6.4.74}$$

then $p^\dagger$ is also symmetric. Thus, we say that $p^\dagger$ is a symmetric extension of p. In fact one can check that $p^\dagger$ is self-adjoint because $D_{p^\dagger} = D_{p^{\dagger\dagger}}$. Thus, we have a self-adjoint extension of p.

A symmetric operator A is essentially self-adjoint if $A^{\dagger\dagger}$ is self-adjoint. What this means is that although A itself is not necessarily self-adjoint there is a unique way to extend it to a self-adjoint operator. That $A^{\dagger\dagger}$ is an extension of A follows from

$$D_A \subset D_{A^\dagger} \subset D_{A^{\dagger\dagger}} \ . \tag{6.4.75}$$

The operator p discussed above is not essentially self-adjoint because, for each value of the parameter θ used to define $D_{p^\dagger}$, we get a different self-adjoint extension. This means that the physical results we get for different values of θ are different. A neat way to say this is that different self-adjoint extensions give different physics. Rather than being a nuisance, this makes the structure of quantum mechanics much richer.

6.5 Cayley Transform: Self-Adjoint Operators

We now examine under what conditions a general symmetric operator possesses self-adjoint extensions and how many. To do this we need some more machinery. The operation analogous to a rotation in a Euclidean space $\mathcal{E}_n$ is a unitary transformation in $\mathcal{H}$. The characteristic property of a rotation in $\mathcal{E}_n$ is that it preserves length and angles or more succinctly, it preserves the inner product. This is also its characteristic in $\mathcal{H}$.

Definition U is unitary iff $D_U = R_U = \mathcal{H}$ and

$$(Uf, Uf) = (f, f) \ . \tag{6.5.76}$$

From this we immediately get

$$U^\dagger U = 1 \ . \tag{6.5.77}$$

And since $D_U = R_U = \mathcal{H}$ we also get

$$UU^\dagger = 1 \ . \tag{6.5.78}$$

Note, unlike the case for finite vector spaces, (6.5.77) does not imply (6.5.78) without the additional assumptions about domains on U. We now show that A is self-adjoint iff the operator

$$U = (A - i1)(A + i1)^{-1} \tag{6.5.79}$$

called the Cayley transform of A is unitary.

Proof

Suppose A is self-adjoint in $\mathcal{H}$ and $f \in D_A$. Then,

$$\begin{aligned} \| Af \pm if \|^2 &= (Af, Af) \pm i(Af, f) \mp i(f, Af) + (f, f) \\ &= \| Af \|^2 + \| f \|^2 \ . \end{aligned} \tag{6.5.80}$$

Therefore, $(A \pm i1)f = 0$ is only possible if $f = 0$. Thus, the operators $(A \pm i1)^{-1}$ and hence U exist. Furthermore, as we now show, the ranges R_{A+i1} and R_{A-i1} are dense in $\mathcal{H}$. For suppose g is orthogonal to all vectors in $R_{A\pm i1}$. Then for $f \in R_{A\pm i1}$ or equivalently, $f = (A \pm i1)h$ we have:

$$0 = (g, f) = (g, Ah \pm ih) = (g, Ah) \pm i(g, h) \ . \tag{6.5.81}$$

Thus,

$$(g, Ah) = \mp i(g, h) \ . \tag{6.5.82}$$

So,

$$g \in D_{A^\dagger} = D_A \tag{6.5.83}$$

and

$$A^\dagger g = Ag = \pm ig \ . \tag{6.5.84}$$

But as we have seen this is not possible unless $g = 0$. Thus, the ranges $R_{A\pm i1}$ are dense in $\mathcal{H}$. We now prove that in fact

$$R_{A\pm i1} = \mathcal{H} \ . \tag{6.5.85}$$

Let $g \in \mathcal{H}$, then since $R_{A\pm i1}$ is dense in $\mathcal{H}$ the limit $g_n = Af_n \pm if_n \to g$ exists. Also, using (6.5.80)

$$\begin{aligned} \| g_n - g_m \|^2 &= \| A(f_n - f_m) \pm i(f_n - f_m) \|^2 \\ &= \| A(f_n - f_m) \|^2 + \| (f_n - f_m) \|^2 \end{aligned} \tag{6.5.86}$$

and thus the f_n and Af_n converge to some vectors f and h respectively. Furthermore, because A is self-adjoint

$$f \in D_A \tag{6.5.87}$$

and

$$h = Af \ . \tag{6.5.88}$$

Hence, by definition of g as the limit of g_n, it follows that

$$g = Af \pm if \in R_{A\pm i1} \ . \tag{6.5.89}$$

Thus, the limit g of the approximating vectors g_n itself belongs to $R_{A\pm i1}$. However, by taking limits of sequences one may obtain any vector in $\mathcal{H}$. Thus, this limit may be any element in $\mathcal{H}$. Hence,

$$R_{A\pm i1} = \mathcal{H} \tag{6.5.90}$$

and we have that

$$D_U = R_U = \mathcal{H} \ . \tag{6.5.91}$$

Now choose any element f, then $f \in D_U$ and hence $f \in D_{(A\pm i1)^{-1}}$. Thus, we can write

$$f = (A + i1)g \tag{6.5.92}$$

and

$$\begin{aligned} Uf &= (A - i1)(A + i1)^{-1}(A + i1)g \\ &= (A - i1)g \ . \end{aligned} \tag{6.5.93}$$

Therefore,

$$\begin{aligned} \| Uf \|^2 &= \| (A - i1)g \|^2 = \| Ag \|^2 + \| g \|^2 \\ &= \| (A + i1)g \|^2 \\ &= \| f \|^2 \ . \end{aligned} \tag{6.5.94}$$

Thus, assuming $A = A^\dagger$ we conclude that $UU^\dagger = U^\dagger U = 1$. It is also always possible to recover A from U according to

$$A = i(1 - U)^{-1}(1 + U) = i(1 + U)(1 - U)^{-1} \ . \tag{6.5.95}$$

We now prove the converse, that if U is unitary, then A is self-adjoint. Let $g \in D_{A^\dagger}$ and define

$$\tilde{g} = A^\dagger g \ . \tag{6.5.96}$$

Then for any $f \in D_A$

$$(g, Af) = (\tilde{g}, f) \ . \tag{6.5.97}$$

But since

$$A = i(1 + U)(1 - U)^{-1} \ , \tag{6.5.98}$$

all $f \in D_A$ are of the form

$$f = (1 - U)h \tag{6.5.99}$$

where

$$h \in D_U = \mathcal{H} \quad . \tag{6.5.100}$$

Therefore, (6.5.97) reads:

$$(g, i(1 + U)h) = (\tilde{g}, (1 - U)h) \tag{6.5.101}$$

for any $h \in \mathcal{H}$. Now since U is unitary and therefore defined everywhere and conserves inner products, we can replace (h, g) by (Uh, Ug) and $(h, \tilde{g})$ by $(Uh, U\tilde{g})$ to get from (6.5.101)

$$(Ug, iUh) + (g, ih) - (U\tilde{g}, Uh) + (\tilde{g}, Uh) = 0 \tag{6.5.102}$$

or

$$(-iUg - ig - U\tilde{g} + \tilde{g}, Uh) = 0 \quad . \tag{6.5.103}$$

Thus, $-iUg - ig - U\tilde{g} + \tilde{g}$ is orthogonal to all elements of $\mathcal{H}$ and hence vanishes

$$-iUg - ig - U\tilde{g} + \tilde{g} = 0 \quad . \tag{6.5.104}$$

From this we get

$$g = -i\tilde{g} - U(g - i\tilde{g}) \; . \tag{6.5.105}$$

We now perform some algebra. Thus,

$$g = \frac{g - i\tilde{g}}{2} - \frac{g + i\tilde{g}}{2} - U(g - i\tilde{g}) \; . \tag{6.5.106}$$

Using (6.5.104) again this becomes

$$g = \frac{g - i\tilde{g}}{2} + \frac{U(g - i\tilde{g})}{2} - U(g - i\tilde{g}) \tag{6.5.107}$$

and hence

$$g = (1 - U)\frac{g - i\tilde{g}}{2} \; . \tag{6.5.108}$$

Similarly we get

$$\tilde{g} = i(1 + U)\frac{g - i\tilde{g}}{2} \; . \tag{6.5.109}$$

This proves two things: If $g \in D_{A^\dagger}$ then a) $g \in D_A$ according to (6.5.108), i.e. it is in $D_{(1-U)^{-1}}$.

$$\text{b) } Ag = \tilde{g} = A^\dagger g \tag{6.5.110}$$

since

$$Ag = i(1+U)(1-U)^{-1}(1-U)\frac{(g-i\tilde{g})}{2} \qquad (6.5.111)$$

or

$$Ag = i(1+U)\frac{(g-i\tilde{g})}{2} = \tilde{g} \; . \qquad (6.5.112)$$

This proves that $A^\dagger = A$ and hence that A is self-adjoint. Before proceeding let us examine the reasons for our interest in self-adjointness. The examination will be, of necessity, somewhat cursory.

6.6 Some Properties of Self-Adjoint Operators

To begin, consider the operator A which means multiplying by a number a. Self-adjointness implies that for $f, g \in \mathcal{H}$

$$(f, ag) = (af, g) \; . \qquad (6.6.113)$$

But according to the definition of the inner product we have

$$(f, ag) = (a^* f, g) \; . \qquad (6.6.114)$$

Thus,

$$a = a^* \; . \qquad (6.6.115)$$

and hence a must be real. This is not a coincidence. In fact self-adjoint operators, in some sense which will become clear, correspond to real numbers. To make this precise we now discuss the eigenvalue problem for self-adjoint operators. Again, let A be a self-adjoint operator; then there are certain vectors belonging to the domain of A on which operations by A are particularly simple. The operation involves multiplication by a number. We have already encountered this in our solution of the Schrödinger equation

$$H\psi_E = E\psi_E \; . \qquad (6.6.116)$$

Here, operating with H on the vector ψ_E involves multiplying ψ_E by E. The vector ψ_E is called an eigenfunction of H belonging to the eigenvalue E. More generally, f_j is an eigenfunction of the operator A belonging to the eigenvalue a_j if

$$Af_j = a_j f_j \; . \qquad (6.6.117)$$

The important properties of self-adjoint operators are that:

a) All eigenvalues of a self-adjoint operator are real.

b) Eigenvectors belonging to different eigenvalues are orthogonal.

c) The eigenvectors form a complete set.

We now prove a) and b). The eigenvalues of a self-adjoint operator are real and the eigenvectors belonging to different eigenvalues are orthogonal.

Proof

Let f_i, f_j be two eigenfunctions of A belonging to the eigenvalues a_i and a_j respectively. Thus,

$$Af_i = a_i f_i \tag{6.6.118}$$

and

$$Af_j = a_j f_j \ . \tag{6.6.119}$$

Forming the inner product of (6.6.118) with f_i we get

$$(f_i, Af_i) = (Af_i, f_i) = (f_i, a_i f_i) = (a_i f_i, f_i) \ . \tag{6.6.120}$$

Therefore,

$$a_i = a_i^* \tag{6.6.121}$$

as required. Note the self-adjointness was necessary for otherwise we do not know that $f_i \in D_{A^\dagger}$. This proves part a).

To prove part b) we use (6.6.118) and (6.6.119) to get

$$(f_i, Af_j) = (Af_i, f_j) \ . \tag{6.6.122}$$

This means that

$$a_j(f_i, f_j) = a_i(f_i, f_j) \tag{6.6.123}$$

so that

$$(a_j - a_i)(f_i, f_j) = 0 \ . \tag{6.6.124}$$

Thus, if $a_i \neq a_j$, then

$$(f_i, f_j) = 0 \ . \tag{6.6.125}$$

Thus, we have established the results.

The proof of completeness of the eigenfunctions of a self-adjoint operator is beyond the scope of this book. Consequently we only show a sort of converse which makes the result appear plausible. The general theorem is known as the *Spectral Theorem* and is discussed in detail in reference [6.1].

Let A be a linear operator with a complete orthonormal set of eigenvectors $\{f_n\}$ and corresponding set of real eigenvalues a_n, then A is self-adjoint. Thus, we have that if

$$Af_n = a_n f_n \tag{6.6.126}$$

and

$$(f_n, f_m) = \delta_{nm} \tag{6.6.127}$$

where the $\{a_n\}$ are real and the $\{f_n\}$ are complete. Then $A = A^\dagger$.

Proof

We must show that $D_A = D_{A^\dagger}$ and that for $f, g \in D_A$

$$(Af, g) = (f, Ag) \; . \tag{6.6.128}$$

The proof is based on knowing A on a basis (the eigenfunctions). Suppose $f, g \in D_A$. Then because the $\{f_n\}$ are complete we have the following expansions

$$\begin{aligned} f &= \sum_n \alpha_n f_n \\ g &= \sum_n \beta_n f_n \end{aligned} \tag{6.6.129}$$

where

$$\begin{aligned} \alpha_n &= (f_n, f) \\ \beta_n &= (f_n, g) \; . \end{aligned} \tag{6.6.130}$$

Suppose $f \in D_A$. Then,

$$\begin{aligned} (f, Af) &= \sum_n (\alpha_m f_m, A\alpha_n f_n) \\ &= \sum_n a_n \alpha_m^* \alpha_n \\ &= \sum_n (a_m \alpha_m f_m, \alpha_n f_n) \\ &= \sum_n (A\alpha_m f_m, \alpha_n f_n) \\ &= (Af, f) = (A^\dagger f, f) \; . \end{aligned} \tag{6.6.131}$$

Therefore, $f \in D_{A^\dagger}$ and in a similar manner we get $(g, Af) = (Ag, f)$. Thus, A is self-adjoint.

6.7 Classification of Symmetric Operators

We now complete the classification of symmetric operators. For the purposes of physics there is no need to distinguish between self-adjoint and essentially self- adjoint operators since the latter always have a unique and obvious extension to self-adjoint operators. We are mainly concerned in determining which symmetric operators have several self-adjoint extensions. Our main tool in this investigation is the Cayley transform that we discussed previously.

Suppose A is an arbitrary symmetric operator (not necessarily self-adjoint). In that case the corresponding Cayley transform U need not be unitary and the domain D_U and the range $R_U = UD_U$ need not coincide with the whole Hilbert space $\mathcal{H}$. If we consider the sets of vectors $D_U^\perp$ and $R_U^\perp$ orthogonal to D_U and R_U, the "size" of these sets gives us an indication of the extent to which U is not unitary and A is not self-adjoint. It is straightforward to check that the sets $D_U^\perp$ and $R_U^\perp$ are in fact subspaces. We call these subspaces the *deficiency subspaces* of A and their dimensions the *deficiency indices.* Thus, the deficiency indices of A are

$$(m, n) = (\dim D_U^\perp\,,\, \dim R_U^\perp)\,. \tag{6.7.132}$$

Now all elements in D_U are of the form

$$f = (A + i1)g\,. \tag{6.7.133}$$

Thus, if $h \in D_U^\perp$, then

$$(h, f) = (h, (A + i1)g) = 0 \tag{6.7.134}$$

or

$$(h, Ag) = -i(h, g) = (ih, g)\,. \tag{6.7.135}$$

So $h \in D_{A^\dagger}$ and

$$A^\dagger h = ih\,. \tag{6.7.136}$$

But, if this is true, then,

$$(h, (A + i1)g) = ((A^\dagger - i1)h, g) = 0 \tag{6.7.137}$$

and $h \in D_U^\perp$. So, we have shown that $h \in D_U^\perp$ iff

$$A^\dagger h = ih \tag{6.7.138}$$

for any $h \in D_{A^\dagger}$.

Thus, $\dim D_U^\perp$ is given by the number of linearly independent solutions of

$$A^\dagger h = ih \tag{6.7.139}$$

belonging to $D_{A^\dagger}$.

Similarly, all vectors in R_U are of the form $(A - i1)g$ and hence by a set of steps like above, we find that $\dim R_U^\perp$ is given by the number of linearly independent solutions of

$$A^\dagger h = -ih \tag{6.7.140}$$

belonging to $D_{A^\dagger}$.

Now from our previous results we know that A is self-adjoint iff U is unitary and hence iff $R_U = D_U = \mathcal{H}$. Thus, A is self-adjoint iff the deficiency indices are $(0, 0)$. To see what this means we return to our previous example of the momentum operator defined on the interval (a,b) and begin with the symmetric, but not self-adjoint operator with domain

$$D_p = \{f \in C^1 \,|\, f(a) = f(b) = 0\} \tag{6.7.141}$$

where C^1 means "functions whose first derivative is continuous". If we now consider the expression

$$\begin{aligned}(h,pf) &= \int_a^b h^* \frac{\hbar}{i}\frac{df}{dx}\,dx \\ &= \frac{\hbar}{i}\,h^* f\big|_a^b + \int_a^b \left(\frac{\hbar}{i}\frac{dh}{dx}\right)^* f\,dx\end{aligned} \tag{6.7.142}$$

it follows from $f(a) = f(b) = 0$ that $h \in D_{p^\dagger}$ for any $h \in \mathcal{H}$ such that both $h, h' \in \mathcal{L}_2$. Furthermore, the action of $p^\dagger$ is the same as that of p, namely $(\hbar/i)\,(d/dx)$.

The only solutions of (6.7.136) and (6.7.140) respectively are:

$$h_+ = Ae^{-x/\hbar} \ , \ h_- = Ae^{x/\hbar} \ . \tag{6.7.143}$$

Thus, the deficiency indices are (1,1).

In general if the deficiency indices of a symmetric operator A are (m,n) with $(m \neq 0, n \neq 0)$ it is possible to extend (increase the domain) the operator A as follows. Let two solutions of (6.7.136) and (6.7.140) be h_+ and h_- respectively. Then for $g \in D_A$ define the operator A' by:

$$A'[g + \theta(h_+ + h_-)] = Ag + i\theta(h_+ - h_-) \ . \tag{6.7.144}$$

Clearly A' is an extension of A since now $h_+ + h_-)$ belongs to $D_{A'}$. Furthermore, the deficiency indices for A' are $(m-1, n-1)$. To see this one simply needs to verify that the corresponding Cayley transform of A, namely U is extended to U' where

$$U'f = Uf \ \text{ if } \ f \in D_U \tag{6.7.145}$$

and

$$U'h_+ = h_- \ . \tag{6.7.146}$$

One can proceed in this manner until one gets deficiency indices $(r,0)$ or $(0,r)$. In this case no further extension is possible. If one then finds that $r \neq 0$, one has to conclude that such an operator has no self-adjoint extensions.

Although the procedure above yields the self-adjoint extensions of symmetric operators with deficiency indices (n,n) it is not the most useful approach. For physicists boundary conditions are usually of important physical significance with direct physical interpretations. This is brought out in the approach we now take. For example, if the deficiency indices are $(1,1)$ we get a one-parameter family of self-adjoint extensions, and for deficiency indices (n,n) we get an n^2-parameter family of self-adjoint extensions. Mathematically this is as far as one can go. To pick the "correct" extension in these cases depends on the physical situation and cannot be decided by mathematics. It requires physics. To illustrate these points we now discuss some specific examples.

Again consider the momentum operator $p = (\hbar/i)\ (d/dx)$ on $\mathcal{L}_2$. Let p be defined on the interval $(-\infty, \infty)$. In this case the equations for the deficiency indices read as before

$$\frac{\hbar}{i}\frac{df}{dx} = \pm i f \tag{6.7.147}$$

or

$$\frac{df}{dx} = \mp\frac{1}{\hbar}f\ . \tag{6.7.148}$$

The solutions are

$$f = Ae^{\mp x/\hbar}\ . \tag{6.7.149}$$

But neither of these solutions is square integrable on $(-\infty, \infty)$ and hence the deficiency indices are $(0, 0)$ and defined over the interval $(-\infty < x < \infty)$ the operator p is self-adjoint.

Now consider the case where the operator $(\hbar/i)\ (d/dx)$ is defined on the interval $(0, \infty)$. As before the solutions are

$$f = Ae^{\mp x/\hbar}\ . \tag{6.7.150}$$

This time

$$f = Ae^{-x/\hbar} \tag{6.7.151}$$

is square integrable, but

$$f = Ae^{x/\hbar} \tag{6.7.152}$$

is not. Thus, the deficiency indices are $(1, 0)$ and defined over $(0 \le x < \infty)$ the operator $(\hbar/i)\,(d/dx)$ has no self-adjoint extensions. The reason for this is easy to see. Consider

$$\left(f, \frac{\hbar}{i}\frac{dg}{dx}\right) = \int_0^\infty (f^*\frac{\hbar}{i}\frac{dg}{dx})\,dx\ . \tag{6.7.153}$$

After integrating by parts, this becomes

$$\begin{aligned}\left(f, \frac{\hbar}{i}\frac{dg}{dx}\right) &= \frac{\hbar}{i}\,f^*g|_0^\infty + \int_0^\infty \left(\frac{\hbar}{i}\frac{df}{dx}\right)^* g\,dx \\ &= \frac{\hbar}{i}\,f^*g|_0^\infty + \left(\frac{\hbar}{i}\frac{df}{dx}, g\right)\ .\end{aligned} \tag{6.7.154}$$

So, for $(\hbar/i)\,(d/dx)$ to be self-adjoint requires that

$$\frac{\hbar}{i}\,f^*g|_0^\infty = 0\ . \tag{6.7.155}$$

Now we know that both f and g vanish for $x \to \infty$. Therefore we require

$$f(0) = 0 \text{ or } g(0) = 0 \,. \tag{6.7.156}$$

In fact to make $p = (\hbar/i)\,(d/dx)$ self-adjoint requires $D_p = D_{p^\dagger}$ and hence both $f(0) = g(0) = 0$. On the other hand, the "eigenfunctions" of this p are Ae^{ikx} and only vanish at $x = 0$ if $A = 0$. Thus, this p would have no eigenfunctions if it were to be self-adjoint. There is also a physical reason for the lack of a self-adjoint extension in this case. We discuss this, in more detail, a little further on.

Finally, consider the operator $p = (\hbar/i)\,(d/dx)$ defined on the interval $a \leq x \leq b$. In this case, both solutions $Ae^{\pm x/\hbar}$ are square integrable and the deficiency indices are $(1,1)$. This means that the self-adjoint extensions should depend on precisely one parameter as we already indicated. We now examine what this parameter is. In this case we have

$$\begin{aligned}(f,pg) &= \int_a^b f^* \frac{\hbar}{i}\frac{dg}{dx}\,dx \\ &= \frac{\hbar}{i}\, f^* g\Big|_a^b + (pf,g) \;.\end{aligned} \tag{6.7.157}$$

Therefore for self-adjointness we require

$$f^*(b)g(b) = f^*(a)g(a) \tag{6.7.158}$$

or

$$\left(\frac{f(b)}{f(a)}\right)^* = \frac{g(a)}{g(b)} \;. \tag{6.7.159}$$

This implies that

$$\frac{f(b)}{f(a)} = \frac{g(b)}{g(a)} = e^{i\theta} \;. \tag{6.7.160}$$

So θ is the parameter determining the different self-adjoint extensions. Actually, this is a specification of boundary conditions. Thus, if we choose $\theta = 0$ we have periodic boundary conditions. These are the most common. Note that once we have specified that the domain of p is the set of all square-integrable functions for $a \leq x \leq b$ (abbreviated $\mathcal{L}_2(a,b)$) such that

$$f(b) = e^{i\theta} f(a) \;, \tag{6.7.161}$$

then the domain of $p^\dagger$ is the same as that of p and the deficiency indices become $(0,0)$ showing that p is self-adjoint. To give some idea of the physics behind these three situations we state a theorem without proof and then use it to explain the physics behind all this.

Stone's Theorem

If A is a self-adjoint operator, then

$$U(s) = e^{isA} \tag{6.7.162}$$

is a unitary operator for every real number s. Furthermore,

$$U(s)U(t) = U(s+t) \tag{6.7.163}$$

and

$$U(-s) = U(s)^{\dagger} = U(s)^{-1} \ . \tag{6.7.164}$$

Conversely given a set of continuous unitary operators satisfying (6.7.163) and (6.7.164) then there is a self-adjoint operator A such that (6.7.162) holds and A is given by

$$iA = \lim_{\epsilon \to 0} \frac{U(\epsilon) - 1}{\epsilon} \ . \tag{6.7.165}$$

Operators satisfying (6.7.163) and (6.7.164) form an algebraic structure called a group. We now construct such a set of unitary operators and use them to interpret what we did.

Consider the operator

$$U(s)f(x) = f(x+s) \ . \tag{6.7.166}$$

Then,

$$U(0)f(x) = f(x) \tag{6.7.167}$$

and

$$U(t)U(s)f(x) = U(t)f(x+s) = f(x+s+t) = U(t+s)f(x) \ . \tag{6.7.168}$$

Thus, the operators defined by (6.7.166) satisfy (6.7.163) and also (6.7.164) as we see by setting $t = -s$ for a left inverse and $s = -t$ for a right inverse. On the other hand, if $f(z)$ is analytic for $a \leq \Re(z) \leq b$ it has the Taylor expansion

$$\begin{aligned} f(x+s) &= \sum_{n}^{\infty} \frac{s^n}{n!} \frac{d^n}{dx^n} f(x) \\ &= \sum_{n}^{\infty} \left(\frac{isp}{\hbar} \right)^n \frac{1}{n!} f(x) \\ &= e^{isp/\hbar} f(x) \ . \end{aligned} \tag{6.7.169}$$

Thus,

$$U(s) = e^{isp/\hbar} \tag{6.7.170}$$

and we have succeeded in expressing $U(s)$ in the form (6.7.162). It makes sense to call $U(s)$ the translation operator since it "translates" functions by an amount s. Furthermore we call p the *generator* of translations since for infinitesimal translations.

$$U(s) \to 1 + i\frac{s}{\hbar}p \quad \text{as} \quad s/\hbar \to 0 \ . \tag{6.7.171}$$

In terms of these considerations we can understand why p has a one-parameter family of self-adjoint extensions on $\mathcal{L}_2(a,b)$. Actually, our considerations depend on the fact that f is analytic for $a \leq \Re(z) \leq b$. This, however, is not a restriction since the functions square-integrable and analytic over $a \leq \Re(z) \leq b$ form a dense set in $\mathcal{L}_2(a,b)$.

Suppose $f(x)$ is an infinitely differentiable function which is non-zero only on an interval completely contained in $a \leq x \leq b$. An example of such a function is

$$f(x) = \begin{cases} 0 & \text{if } x \leq \alpha < a \ , \ x \geq \beta > b \\ \exp\left(-\frac{1}{x-\alpha} - \frac{1}{\beta - x}\right) & \text{if } \alpha \leq x \leq \beta \end{cases} \quad . \tag{6.7.172}$$

Note that this function is definitely not analytic over $a \leq x \leq b$ since it has essential singularities at both $x = \alpha$ and $x = \beta$. Also, an analytic function cannot vanish on an open interval unless it is identically zero. Now consider the normalization (i.e. unit probability) associated with f, namely, $\int_a^b |f(x)|^2 \, dx$. We want the translation operator $U(s)$ to preserve this normalization, that is, to be unitary. Thus, we need

$$\int_a^b |U(s)f(x)|^2 \, dx = \int_a^b |f(x)|^2 \, dx \ . \tag{6.7.173}$$

But,

$$U(s)f(x) = f(x+s) \tag{6.7.174}$$

and if $s > b - \beta$ part of the wavefunction "disappears" past the right end point. To conserve the integral above requires that what disappears at the right must reappear from the left. Of course the phase of the function can be shifted in reappearing from the left. Furthermore, all functions must experience the same phase shift. Thus, if f_1 and f_2 are two such functions and if their phase shifts are different, say θ_1 and θ_2, then translation of the function $f = f_1 + f_2$ will eventually produce the function $f' = e^{i\theta_1} f_1 + e^{i\theta_2} f_2$. But,

$$\int_a^b |U(s)f(x)|^2 \, dx \neq \int_a^b |f(x)|^2 \, dx \tag{6.7.175}$$

unless $\theta_1 = \theta_2$. So the superposition principle limits the number of phase shift parameters to one.

Why then does p not have any self-adjoint extensions on $\mathcal{L}_2(0,\infty)$? The answer is as follows. Translating x to the right will never take the function past the right endpoint. On the other hand, by translating to the left we can always bring the function past the left endpoint (the origin). In this case there is not anywhere from where the function can reappear to conserve probability and hence p can not be made self-adjoint. This also explains why p is already self-adjoint on $\mathcal{L}_2(-\infty,\infty)$.

Another extremely simple problem is the case of a particle in a strongly repulsive potential such as a quartic or cubic potential

$$V(x) = -gx^n \quad g > 0 \quad n = 3, 4 \ . \tag{6.7.176}$$

In this case the Hamiltonian is

$$H = \frac{p^2}{2m} + V(x) = \frac{p^2}{2m} - gx^n \,. \tag{6.7.177}$$

By redefining the variables we can bring this to the form

$$H = -\frac{d^2}{dx^2} - ax^n \quad a > 0 \,. \tag{6.7.178}$$

We want to examine this Hamiltonian on $\mathcal{L}_2(-\infty, \infty)$. To get a feel for the physical situation consider the problem classically. In the previous example we saw that the existence of different self-adjoint extensions depended on the fact that the particle can reach a boundary (a or b) and have to be transmitted or reflected.

In this case the boundaries are at $\pm\infty$. So we must see if the particle can in fact reach these boundaries. Now suppose the particle starts at $x = 0$ with energy $E > 0$. Then classically its velocity v is given by

$$\frac{1}{2}mv^2 - gx^n = E \tag{6.7.179}$$

or

$$v = \sqrt{\frac{2E}{m} + \frac{2}{m}gx^n} \,. \tag{6.7.180}$$

Therefore the time to reach ∞ is

$$t = \int_0^\infty \frac{dx}{v} = \int_0^\infty \frac{dx}{\sqrt{\frac{2E}{m} + \frac{2}{m}gx^n}} \,. \tag{6.7.181}$$

And for $n = 3$ or 4, $t < \infty$. So, the particle reaches $+\infty$ in a finite time. To conserve probability it must be reflected and return to the origin in a finite time. Thus, the time-averaged particle position is near the origin. We therefore expect to find that all eigenfunctions of this Hamiltonian are square integrable and that the spectra of the self-adjoint extensions of H are discrete. This is, in fact, the case and this Hamiltonian is analogous to a free particle Hamiltonian on a finite finite interval (a, b), (see problem 6.7). For the repulsive quartic potential the points $\pm\infty$ behave like the end points (a, b). So it is not surprising that the deficiency indices are $(2, 2)$ for both cases. For the repulsive cubic potential the situation is different. The particle can again reach $x = +\infty$ in a finite time but it can never reach $x = -\infty$. Thus we need only specify boundary conditions at $x = +\infty$. In this case the deficiency indices turn out to be $(1, 1)$. The main point of this discussion is that whenever an operator, which is a candidate for representing an observable, is not self-adjoint but has self-adjoint extensions, then there are good physical reasons for this.

6.8 Spontaneously Broken Symmetry

The concept of spontaneously broken symmetries plays a very important role in some field theories of elementary particles. Since it fits naturally into the topics we have just discussed, we shall start by defining the concept and then proceed to illustrate it with an example.

Suppose we have some observable, whose representative operator Q commutes with the Hamiltonian H. Then either Q corresponds to a discrete symmetry operation Q_D such as parity, or else we can use Q to define a one-parameter family of unitary operators,

$$U(\alpha) = e^{i\alpha Q} \cdot \tag{6.8.182}$$

for which Q is the generator. It then follows from

$$[Q, H] = 0 \tag{6.8.183}$$

that

$$e^{i\alpha Q} H e^{-i\alpha Q} = H \tag{6.8.184}$$

or else for a discrete symmetry

$$Q_D H Q_D^{-1} = H\ . \tag{6.8.185}$$

We also require that the ground state of the Hamiltonian ϕ_0 should be invariant under either $U(\alpha)$ or Q_D. This means

$$U(\alpha)\phi_0 = \phi_0 \tag{6.8.186}$$

or

$$Q_D\phi_0 = \phi_0\ \ . \tag{6.8.187}$$

Definition

A symmetry corresponding to an observable Q (Q_D) is spontaneously broken if all the above statements except (6.8.186) (respectively (6.8.187)) hold. For this to occur requires that the ground state be degenerate. This, in itself, is an unusual phenomenon. To illustrate this phenomenon we consider the Hamiltonian

$$H = \frac{p^2}{2m} \tag{6.8.188}$$

defined on the interval $-a \leq x \leq a$. We furthermore pick for $p = (\hbar/i)\ (d/dx)$ the self-adjoint extension corresponding to the domain

$$D_p = \left\{ f(x) \in C^1 \mid f(a) = -f(-a) \right\}\ . \tag{6.8.189}$$

Thus, instead of periodic we pick "anti-periodic" boundary conditions.

The complete set of normalized eigenfunctions of this momentum operator are given by:

$$f_n(x) = \frac{1}{\sqrt{2a}} e^{i\pi(n+1/2)x/a} \quad n = 0, \pm 1, \pm 2, \ldots \tag{6.8.190}$$

with corresponding eigenvalues $(n+1/2)\pi\hbar/a$. These wavefunctions have the following symmetry properties

$$f_n(x) = f_{-(n+1)}(-x) \tag{6.8.191}$$

and

$$f_n^*(x) = f_{-(n+1)}(x) \ . \tag{6.8.192}$$

Thus, the parity operator P and the time-reversal operator T have the following action on them

$$(Pf_n)(x) = f_{-(n+1)}(x) \tag{6.8.193}$$

$$T(f_n)(x) = f_{-(n+1)}(x) \ . \tag{6.8.194}$$

The set of functions $\{f_n\}$ are also eigenfunctions of the Hamiltonian. In fact,

$$Hf_n = \frac{\pi^2\hbar^2}{2ma^2}(n+1/2)^2 f_n \tag{6.8.195}$$

$$Hf_{-(n+1)} = \frac{\pi^2\hbar^2}{2ma^2}(n+1/2)^2 f_{-(n+1)} \ . \tag{6.8.196}$$

Thus, all eigenvalues including the ground state eigenvalue

$$E_0 = E_{-1} = (\pi^2\hbar^2)/(2ma^2)$$

are *doubly degenerate*. We further see that although the Hamiltonian H, the parity operator P, and the time-reversal operator T commute, the two ground states f_0 and f_{-1} are not eigenstates of either the parity operator or the time-reversal operator

$$(Pf_0)(x) = T(f_0)(x) = f_{-1}(x) \tag{6.8.197}$$

and

$$(Pf_{-1})(x) = T(f_{-1})(x) = f_0(x) \ . \tag{6.8.198}$$

Thus, parity and time-reversal are spontaneously broken symmetries.

It is possible to restore these symmetries by defining states

$$g_n^+(x) = \frac{1}{\sqrt{2}}\left[f_n(x) + f_{-(n+1)}(x)\right] = \frac{1}{\sqrt{a}}\cos(n+1/2)\frac{\pi x}{a} \tag{6.8.199}$$

as well as

$$g_n^-(x) = \frac{1}{\sqrt{2}}\left[-f_n(x) + f_{-(n+1)}(x)\right] = \frac{1}{\sqrt{a}}\sin(n+1/2)\frac{\pi x}{a} \ . \tag{6.8.200}$$

These are now simultaneous eigenstates of H, P, and T.

$$(Pg_n)^\pm(x) = \pm\, g_n^\pm(x) \tag{6.8.201}$$

$$(Tg_n)^\pm(x) = \pm\, g_n^\pm(x) \tag{6.8.202}$$

$$Hg_n^\pm(x) = \frac{\pi^2\hbar^2}{2ma^2}\, g_n^\pm(x) \ . \tag{6.8.203}$$

So, parity and time reversal are no longer broken symmetries. In this case, however, we have an even more surprising symmetry breaking, for although the momentum operator p and the Hamiltonian $p^2/2m$ commute, these eigenstates of the Hamiltonian are not eigenstates of the momentum operator. In fact,

$$pg_0^\pm(x) = \mp\frac{\pi\hbar}{a}(n+1/2)\, g_n^\mp(x) \tag{6.8.204}$$

so that in particular

$$pg_0^+(x) = -\frac{\pi\hbar}{2a}\, g_0^-(x) \tag{6.8.205}$$

and

$$pg_0^-(x) = +\frac{\pi\hbar}{2a}\, g_0^+(x) \ . \tag{6.8.206}$$

In this case we therefore have translational symmetry spontaneously broken since the translation operator

$$U(s) = e^{isp/\hbar} \tag{6.8.207}$$

does not leave the ground states $g_0^\pm$ invariant. In fact, by expanding

$$U(s) = \sum_n \frac{1}{n!}\left(\frac{isp}{\hbar}\right)^n \tag{6.8.208}$$

and repeatedly applying (6.8.205) and (6.8.206) we get:

$$U(s)g_0^\pm(x) = \cos\left(\frac{\pi s}{2a}\right)\, g_0^\pm(x) \mp i\sin\left(\frac{\pi s}{2a}\right)\, g_0^\mp(x) \ . \tag{6.8.209}$$

This demonstrates conclusively that the translational symmetry is broken.

We now relate the mathematical model we have displayed, to a definite physical system. If one considers a one-dimensional crystal consisting of only one type of atom, then the boundary condition in going from nearest neighbour to nearest neighbour is periodic. The situation repeats itself. Similarly for a one-dimensional crystal with alternating atoms $(ABAB)\ldots$, as in an antiferromagnet, the boundary condition from an atom to its next nearest neighbour is periodic, and hence from nearest neighbour to nearest neighbour anti-periodic.

We can now visualize the physical situation corresponding to our model and get a clearer understanding of the cause of the broken symmetry. If we consider such an antiferromagnetic crystal and consider the interval between nearest neighbours as fundamental, we must impose antiperiodic boundary conditions. Furthermore, since the end points correspond physically to different situations (atoms) it makes a difference whether a particle travels freely from left to right or right to left. The situations are not mirror images of each other and hence not eigenstates of the parity operator. Since time-reversal reverses the direction of travel, these states are also not eigenstates of the time-reversal operator.

One can take a superposition of states of particles travelling to the left and right, as we did, to get standing waves which are then automatically time-reversal as well as parity invariant. In this case, however, conservation of probability brings about a loss of translation invariance. It is clear now that this "unusual" self-adjoint extension of the momentum operator has just as physical an interpretation as the usual one with periodic boundary conditions.

It is perhaps also worth while to notice that the commutation relation

$$[x,p] = i\hbar \tag{6.8.210}$$

is not valid in this representation since for $f \in D_p$, $xf \notin D_p$ in general. In fact, in this case, $xf \in D_p$ only if $f(a) = f(-a) = 0$. Nevertheless, it is true that

$$[x^{2n},p] = i2n\hbar\, x^{2n-1} \quad n = 0,1,2,\dots\ . \tag{6.8.211}$$

This concludes our mathematical treatment of self-adjointness. We now turn to a systematic analysis of the physical interpretation of quantum mechanics.

6.9 Problems

6.1 Consider the set of functions $\{f_k(x) = e^{ikx}f(x)\ ,\ f(x) \in \mathcal{L}_2\}$. Show that

$$\lim_{k\to\infty}(g,f_k) = 0 \quad g \in \mathcal{L}_2$$

whereas

$$\| f_k \|^2 = \| f \|^2 \neq 0\ .$$

The above type of convergence of $f_k \to 0$ is called weak convergence as opposed to the notion of strong convergence defined in the text. Hint: Use the Riemann-Lebesgue Theorem [1].

6.2 Show that every Cauchy sequence in a finite dimensional vector space converges strongly (see problem 6.1).

6.3 Consider the operator $A = px^{2n+1} + x^{2n+1}p$ where $p = (\hbar/i)(d/dx)$ and $n = 1,2,3,\dots$. Find the eigenvalues and eigenfunctions of A. What are the deficiency indices of A ? The Hilbert space in this case is $\mathcal{L}_2(-\infty,\infty)$. For $n = 1$ this example is due to von Neumann.

6.4 A projection operator is a self-adjoint, non-negative operator P satisfying $P^2 = P$. Let f_n be a normalized eigenfunction of a self-adjoint operator A with only discrete eigenvalues λ_n .
a) Show that the operator $P_n\psi = f_n(f_n, \psi)$ is a projection operator.
b) Show that A can be written

$$A\psi = \sum_n \int \lambda_n f_n(x) f_n^*(y)\psi(y)\,dy = \sum_n \lambda_n P_n \psi \ .$$

This is called the spectral resolution of the operator A.
Hint: Assume completeness of the eigenfunctions.

6.5 Find the spectral resolution (see problem 6.4) of the operator

$$A = \begin{pmatrix} a_3 & a_1 + ia_2 \\ a_1 - ia_2 & -a_3 \end{pmatrix} \quad a_i \text{ real.}$$

6.6 For any operator A the corresponding operator $R(z) = (A - z1)^{-1}$, where 1 stands for the unit operator, is called the resolvent operator. Show that for any square matrix A , $R(z)$ is analytic in z with poles at the eigenvalues of A.

6.7 Find the deficiency indices and hence all self-adjoint extensions of the Hamiltonian

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$$

defined on the interval (a, b).
Hint: It may be useful to express the boundary conditions on a function $f \in D_H$ in terms of 2-component quantities

$$F(a) = \begin{pmatrix} f(a) \\ f'(a) \end{pmatrix} \quad \text{and} \quad F(b) = \begin{pmatrix} f(b) \\ f'(b) \end{pmatrix}$$

and assume that $F(b) = UF(a)$ where U is a non-singular 2×2 matrix.

6.8 Given an orthonormal basis set $\{u_n : n = 0, 1, 2, \ldots\}$ and an operator a which has the following action on this basis:

$$au_n = \sqrt{n}\, u_{n-1} \quad n \geq 0 \ .$$

Find the adjoint operator $a^\dagger$ by explicitly giving its action on this basis set. Also find the commutator $[a, a^\dagger]$.

Bibliography

[6.1] G.F. Simmons, *Introduction to Topology and Modern Analysis* - McGraw-Hill Book Co. (1963) Chapter 10.

[6.2] A.S. Wightman, Chapter 8 in *Cargese Lectures in Theoretical Physics*, Edited by M. Levy - Gordon and Breach Inc. (1967).

[6.3] A.H. Zemanian, *Distribution Theory and Transform Analysis* - McGraw-Hill Book Co. (1965) page 174.

[6.4] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* - Princeton University Press (1955).

[6.5] T. Kato, *Perturbation Theory for Linear Operators* - Springer-Verlag, New York Inc. (1966). This book gives a very concise and complete summary of useful Hilbert space techniques.

[6.6] Two of the standard treatises on functional analysis are:
F. Riesz and B.S.Z. Nagy, *Functional Analysis* - Frederick Unger Publishing Co., N.Y. (1955) and
K. Yosida - *Functional Analysis* - Springer-Verlag, New York Inc. (1966).

Chapter 7

Physical Interpretation

7.1 Introduction

At this stage we have developed all the formal mathematical machinery that we need. This does not mean we have developed all the techniques needed or useful in solving concrete problems. What we have developed is a machinery that allows us to form the physical interpretation of the theory in a precise and economical manner. We do this by displaying classical and quantum mechanics side by side.

In classical mechanics the state of a physical system is described by a point in phase space. Thus, specifying x and p for each one of a system of point particles specifies the system completely. For simplicity we consider systems consisting of a single point particle. Thus, in classical mechanics there is no distinction between specifying the values of certain observables (x, p) and specifying the state of a system. In quantum mechanics the situation is quite different. Here, there is a definite distinction. We now state this in terms of a series of assumptions or axioms.

7.2 A1 - Physical States

The state of a physical system is completely specified by a ray in Hilbert space. A ray is a constant multiple of a vector. Conversely to every vector in Hilbert space corresponds the state of a physical system.

This last assumption needs to be modified in certain cases, called superselection rules, with which we are not concerned. Also, these modifications do not occur for the physics we discuss so we ignore them. Notice, that A1 is a very strong assumption in that it states that the physical state is completely specified by the Hilbert space ray. Thus, if $f \in \mathcal{H}$, then f and cf (where c is a complex number) both describe the same state and furthermore, everything that can in principle be determined about the state is contained in f. Furthermore, A1 says nothing about observables. This is contained in our next assumption, but first we must define observables. *In classical mechanics an observable is any dynamical variable or any function of dynamical variables.* In quantum mechanics the situation is again quite different. We define observables as follows.

In quantum mechanics, an observable is *any physical quantity whose value is obtained by a definite physical operation.* Thus, the physical operation or method of measuring defines the observable. The measurement need not be performed but it must be possible, in principle, to perform the operation yielding the measurement. Furthermore, we shall assume that the measurement operation is ideal in the sense that all experimental errors are zero. It will also turn out that there is not a one-one correspondence between observables in classical mechanics and observables in quantum mechanics To be explicit we shall denote observables by script letters $\mathcal{A}, \mathcal{B}, \mathcal{C}$ etc. Another point is worth mentioning. In classical mechanics we do not distinguish between the mathematical representation of an observable and the value of an observable. Thus $x(t)$ represents both the function and the value of the position at time t. In quantum mechanics the situation is again quite different.

7.3 A2 - Observables

Any physical observable $\mathcal{A}$ is represented in quantum mechanics by a self-adjoint operator A in the Hilbert space of physical states. Furthermore, since A is self-adjoint it possesses a complete set of eigenfunctions with a corresponding set of real eigenvalues

$$A\, f_j = a_j\, f_j \quad . \tag{7.3.1}$$

Also any measurement of $\mathcal{A}$ can yield as a value only one of the eigenvalues a_j and no other number. Conversely to every self-adjoint operator A there corresponds a physical observable $\mathcal{A}$. For simplicity we assume that all the eigenvalues a_j are distinct. In that case the eigenfunctions f_j are orthogonal and we may as well consider them normalized. Thus, the $\{f_j\}$ form a basis set in $\mathcal{H}$ called the eigenbasis of A. This means that any physical state f can be written as a linear superposition of the f_j

$$f = \sum (f_j, f) f_j \; . \tag{7.3.2}$$

This has important implications. Notice also that A being self-adjoint guarantees that the value of $\mathcal{A}$ is a real number, namely one of the a_j . We now know that the possible outcome of a measurement is one of the a_j but we do not know which one or, if this cannot be stated, what is the probability for a given one. This is in fact the content of our next assumption.

7.4 A3 - Probabilities

For an observable $\mathcal{A}$ as specified in (A2) and any physical state f (assumed normalized) the most detailed statement one can make regarding a measurement of $\mathcal{A}$ is that any one of the eigenvalues of A may occur and that the probability P_j that a given eigenvalue a_j occurs is given by:

$$P_j = |(f_j, f)|^2 \,. \tag{7.4.3}$$

Notice that (f_j, f) is the coefficient of f_j in the expansion of f as a linear superposition of the $\{f_j\}$. Thus, writing

$$f = \sum \alpha_j f_j \tag{7.4.4}$$

we have

$$\alpha_j = (f_j, f) \,. \tag{7.4.5}$$

It is therefore appropriate to call α_j the *probability amplitude* for observing the value a_j since this probability is given by

$$P_j = |\alpha_j|^2 \,. \tag{7.4.6}$$

This is of course the point at which quantum mechanics differs most radically from classical mechanics. Thus, although we may know a state completely, we are nevertheless unable, in general, to predict the outcome of an experiment to measure a given observable with certainty. Even worse if we make two separate but identical measurements on a system taking due care that in both cases the system is in exactly the same state before each measurement, the results of the two measurements will generally differ. The point is that states and observables are defined differently in quantum mechanics whereas a state in classical mechanics is defined in terms of observables.

We must now check that our interpretation is consistent. This requires some elementary considerations. In order that P_j be a probability we need

1) $0 \leq P_j \leq 1$

2) $\sum_j P_j = 1$

The first follows trivially from Schwarz's inequality since

$$0 \leq |(f_j, f)|^2 \leq \| f_j \|^2 \| f \|^2 = 1 \,. \tag{7.4.7}$$

The second follows from (7.3.2)

$$f = \sum (f_j, f) f_j \,. \tag{7.4.8}$$

Therefore,

$$(f, f) = 1 = \sum (f_j, f)(f, f_j) = \sum |(f_j, f)|^2 \,. \tag{7.4.9}$$

Although in general we can only make predictions of probabilities, there are occasions when predictions can be made with absolute certainty. For example, if in (7.4.6)

$$\alpha_j = 0 \quad j \neq k \tag{7.4.10}$$

and

$$\alpha_j = 1 \quad j = k \tag{7.4.11}$$

we get that

$$P_k = 1 \quad \text{and} \quad P_j = 0 \quad j \neq k\,. \tag{7.4.12}$$

Thus, in this case the result of a measurement is certain to be a_k . Furthermore, using (7.4.4) we see that the state on which the measurement is performed is $f = f_k$.

Thus, if the system is in an eigenstate of an observable, then a measurement of that observable is certain to yield the eigenvalue corresponding to this eigenstate. This in no way implies that the states which are eigenstates of the operator A corresponding to some observable $\mathcal{A}$ are more precisely specified than states which are not eigenstates of A, because if $\mathcal{B}$ is another observable, the eigenstates of A (the operator corresponding to $\mathcal{A}$) are generally not also eigenstates of B (the operator corresponding to $\mathcal{B}$). Thus, although the value of $\mathcal{A}$ will be well-defined in this state the value of $\mathcal{B}$ will not. In fact we shall agree that *an observable has a value only for states which are eigenstates of this observable.* Of course if the state under consideration is "almost" an eigenstate then we may say that the observable has approximately the value corresponding to the predominant eigenstate. To make this more precise we introduce the concepts of average or expectation value of an observable and the root-mean-square (RMS) deviations of an observable.

The expectation or average value $\langle A\rangle$ of an observable $\mathcal{A}$ is given by summing all possible values multiplied by their probabilities. Thus, using assumption A2 we get:

$$\langle A\rangle = \sum_j P_j\, a_j \tag{7.4.13}$$

and using (7.4.3) this becomes

$$\begin{aligned}\langle A\rangle &= \sum_j |(f_j, f)|^2\, a_j \\ &= \sum_j (f, f_j)(f_j, f)\, a_j \\ &= \sum_j (f, Af_j)(f_j, f)\,.\end{aligned} \tag{7.4.14}$$

Since A is self-adjoint, and therefore has a complete set of eigenstates, we can write this

$$\begin{aligned}\langle A\rangle &= \sum_j (f, Af_j)(f_j, f) \\ &= \sum_j (Af, f_j)(f_j, f) \\ &= (Af, f) = (f, Af)\,.\end{aligned} \tag{7.4.15}$$

Thus,

$$\langle A\rangle = (f, Af) \ . \tag{7.4.16}$$

The RMS deviation ΔA is defined by:

$$\begin{aligned}(\Delta A)^2 &= \langle (A - \langle A\rangle)^2\rangle \\ &= \langle A^2 - 2A\langle A\rangle + \langle A\rangle^2\rangle \\ &= \langle A^2\rangle - 2\langle A\rangle\langle A\rangle + \langle A\rangle^2 \ .\end{aligned} \tag{7.4.17}$$

Therefore,

$$(\Delta A)^2 = \langle A^2\rangle - \langle A\rangle^2 \ . \tag{7.4.18}$$

The RMS value ΔA provides a measure of how spread out the "value" of $\mathcal{A}$ is. Thus, if f is an eigenstate say f_j. Then,

$$\begin{aligned}\langle A\rangle &= (f_j, Af_j) \\ &= (f_j, a_j f_j) \\ &= a_j(f_j, f_j) = a_j \ .\end{aligned} \tag{7.4.19}$$

Also, in this case, it is obvious from the definition that

$$\Delta A = 0 \tag{7.4.20}$$

since,

$$\begin{aligned}(\Delta A)^2 &= \langle A^2\rangle - \langle A\rangle^2 \\ &= \langle f_j, A^2 f_j\rangle - a_j^2 \\ &= a_j^2 - a_j^2 = 0 \ .\end{aligned} \tag{7.4.21}$$

Thus, in an eigenstate the RMS deviation from the expectation value is zero. This means that $\mathcal{A}$ has a sharp value, and in this case the expectation value will coincide with the observed value. Thus, the expectation value corresponds to the average of a large number of identical measurements on systems in the same state prior to the measurement. It may be worth pointing out that the expectation value will not usually coincide with any of the actually measured values.

So far we have talked about performing measurements on systems in the same state. This implies that we have somehow "prepared" the systems in this state. We now show how to do this. For although two separate measurements of an observable made on systems in the same state do not generally yield the same result, we do not have complete unpredictability. In fact if we make a measurement on a system and then immediately repeat it before the system has a chance to evolve, the results of the two measurements are identical. This is in fact the content of our next assumption.

7.5 A4 - Reduction of the Wave Packet

If the system is in a state specified by the wavefunction ψ and a measurement of an observable $\mathcal{A}$ is made yielding the value a_j , then immediately after the measurement the state of the system will be specified by the wavefunction f_j where f_j is an eigenfunction of the operator A corresponding to the eigenvalue a_j. Also A is the operator corresponding to the observable $\mathcal{A}$. Thus,

$$Af_j = a_j f_j \; . \tag{7.5.22}$$

The wording above was deliberate when we stated that f_j is an eigenfunction of A. This is because it is possible for an operator to have several eigenfunctions corresponding to one eigenvalue. In that case we say that the eigenvalue is degenerate and any linear combination of the eigenfunctions corresponding to this eigenvalue is also an eigenfunction. Thus, the wavefunction f_j in the assumption above may be any such linear combination if a_j is degenerate. Suppose a_j is non-degenerate, then we know the state precisely and so to perform a measurement in quantum mechanics is the same as to prepare a state. In fact, that is how we prepare states quantum mechanically. Thus, unless we are in an eigenstate of A it is impossible to predict, with certainty, the result of a measurement of $\mathcal{A}$ prior to the measurement. Performing the measurement causes an uncontrollable change in the wavefunction. This means that a single measurement cannot tell us anything about what the state of the system was prior to the measurement, it only tells us what the state is immediately after the measurement. This is quite different from what a measurement does in classical mechanics. In classical mechanics an observable always has a value and in principle we can measure its value without disturbing the system. Thus, in classical mechanics a measurement tells us both what the state of the system was before as well as what the state of the system is after the measurement. In quantum mechanics we can only determine what the state is after the measurement. There is *no retrodiction*, no specifying of what the state was before the measurement.

7.5.1 Example

Consider the operator

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.5.23}$$

corresponding to an "observable" in a two-dimensional Hilbert space. Clearly the eigenvalues are ± 1 and the corresponding eigenvectors are

$$f_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad f_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \; . \tag{7.5.24}$$

Suppose we are given a wave-function

$$\psi = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \; . \tag{7.5.25}$$

Then,

$$(\psi,\psi) = \sin^2\theta + \cos^2\theta = 1\ . \tag{7.5.26}$$

So ψ is normalized. Also,

$$\psi = \cos\theta f_+ + \sin\theta f_-\ . \tag{7.5.27}$$

A measurement of the observable corresponding to σ_3 must, therefore, yield either $+1$ or -1. The corresponding probabilities are $\cos^2\theta$ and $\sin^2\theta$. The expectation value is

$$\langle\sigma_3\rangle = (\psi,\sigma_3\psi) = \cos^2\theta - \sin^2\theta = \cos 2\theta \tag{7.5.28}$$

and can lie anywhere between -1 and $+1$ depending on the value of θ. Thus, if $\theta = 0$ or $\psi = f_+$, then

$$\langle\sigma_3\rangle = +1\ . \tag{7.5.29}$$

If $\theta = \pi/2$ or $\psi = f_-$ then,

$$\langle\sigma_3\rangle = -1\ . \tag{7.5.30}$$

The RMS deviation is given by:

$$(\Delta\sigma_3)^2 = \langle\sigma_3^2\rangle - \langle\sigma_3\rangle^2\ . \tag{7.5.31}$$

But,

$$(\sigma_3)^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{7.5.32}$$

so

$$\langle(\sigma_3)^2\rangle = +1\ . \tag{7.5.33}$$

Therefore,

$$(\Delta\sigma_3)^2 = 1 - \cos^2\theta = \sin^2\theta\ . \tag{7.5.34}$$

Thus, for $\theta = 0$ or $\pi/2$

$$(\Delta\sigma_3)^2 = 0\ . \tag{7.5.35}$$

This means that if ψ is an eigenstate, $(\Delta\sigma_3)^2 = 0$. On the other hand, for any value $0 < \theta < \pi/2$, σ_3 does not have a definite value but only a probability for obtaining $+1$ or -1 can be given. For example if $\theta = \pi/4$ we have a 50% chance of obtaining either $+1$ or -1 in a measurement. Incidentally σ_3 does correspond to a physical observable and we shall encounter this operator again later.

So far we have only considered measurements on a single observable and have found a profound difference between classical mechanics and quantum mechanics We now consider measurements of several *different observables*. This will tend to accentuate the difference between classical mechanics and quantum mechanics.

As stated before, in classical mechanics every dynamical variable has a definite value in every conceivable state of the system. On the other hand, in quantum mechanics a dynamical variable or observable has a definite value only if the system is in an eigenstate of the corresponding operator. If we now consider two observables, they can both have sharp values iff the system is a simultaneous eigenstate of both of them. What does this entail?

Let $\mathcal{A}$ and $\mathcal{B}$ be two observables with the corresponding self-adjoint operators A and B having a *common dense domain.* Then if a measurement of $\mathcal{A}$ yielding a_n, is immediately followed by a measurement of $\mathcal{B}$ and a second measurement of $\mathcal{A}$, we say that $\mathcal{A}$ and $\mathcal{B}$ are compatible if the second value obtained for $\mathcal{A}$ is always the same as the first (namely a_n in this case). We saw an example of compatible operators in the case of an Hamiltonian H with a potential that is an even function of x, namely $V(x) = V(-x)$. In that case the operators for corresponding compatible observables were the Hamiltonian H and the parity operator P. In fact the following three statements are equivalent.

7.6 Compatibility Theorem and Uncertainty Principle

The following statements are equivalent for a pair of observables $\mathcal{A}$, $\mathcal{B}$.
1) $\mathcal{A}$ and $\mathcal{B}$ are compatible.
2) A and B possess a common eigenbasis.
3) A and B commute, that is

$$[A,B] = AB - BA = 0 \,.$$

An example of two observables that are always compatible is $\mathcal{A}$ and $f(\mathcal{A})$, any function of $\mathcal{A}$. This is reasonable in view of how the operator $f(A)$ corresponding to $f(\mathcal{A})$ is obtained (as a power series in A).

We are now ready to discuss the *Heisenberg Uncertainty Principle.* Suppose $\mathcal{A}$ and $\mathcal{B}$ are not compatible, then $[A,B] \neq 0$. This means that there exists at least one state f belonging to the common domain of A and B such that

$$(AB - BA)f \neq 0 \,. \tag{7.6.36}$$

Now there are many pairs of observables (in fact all classically conjugate variables) for which the commutator is particularly simple, namely just a number. For example, using

$$p = \frac{\hbar}{i}\,\frac{d}{dx} \tag{7.6.37}$$

we have

$$[p,x]f = \frac{\hbar}{i}\,\frac{d}{dx}(xf) - \frac{\hbar}{i}\,x\,\frac{df}{dx} = -i\hbar\, f \tag{7.6.38}$$

so that

$$[p,x]f = -i\hbar\, f \,. \tag{7.6.39}$$

For such observables their lack of compatibility is expressed by the Heisenberg uncertainty principle which states:

Heisenberg Uncertainty Principle

Let the state of a system be described by the wavefunction ψ and let $\mathcal{A}$ and $\mathcal{B}$ be two observables. Then the uncertainties (RMS deviations from the mean) in $\mathcal{A}$ and $\mathcal{B}$ satisfy the inequality

$$\Delta A \Delta B \geq \frac{1}{2} \left|(\psi, [A, B]\psi)\right| . \tag{7.6.40}$$

This is a rather formidable looking expression but if $[A, B]$ is just a c-number (ordinary number rather than an operator which is called a q-number), then the expression simplifies. Thus if

$$[A, B] = c \tag{7.6.41}$$

then (7.6.40) reads, for a normalized ψ,

$$\Delta A \Delta B \geq \frac{1}{2}|c| . \tag{7.6.42}$$

We shall derive (7.6.40) shortly but first let us consider its implications. Classically observables are not operators and therefore the commutator is always 0, implying that all observables are compatible. Now for conjugate variables such as p and x in the example above, $c = i\hbar$ so that

$$\Delta p \Delta x \geq \frac{\hbar}{2} . \tag{7.6.43}$$

Thus, we see that in some sense classical mechanics corresponds to the limit $\hbar = 0$. Of course we know this already from the way Planck introduced h as the smallest lump of action possible.

In quantum mechanics according to (7.6.42) we see, however, that if we increase the precision with which a given observable is known, the corresponding non-compatible observable becomes more uncertain. As we shall see, for p and x this is strictly a property of Fourier transforms. However, it is one of the most outstanding features of quantum mechanics. We now give a physical derivation of (7.6.43) using the so-called Heisenberg microscope and then we proceed to the mathematical derivation.

7.7 The Heisenberg Microscope

The experiment considered here is not an actual but rather a "Gedanken" or thought experiment: We have a microscope as shown (figure 7.1) illuminated by light from the left. Suppose we are examining an electron represented by the little sphere. Also assume that the electron is at rest. We now try to determine its position as precisely as possible. We already know that the momentum is zero. Now in order to "observe" the electron, we must scatter at least one photon off it. This will impart some momentum to it and since the aperture

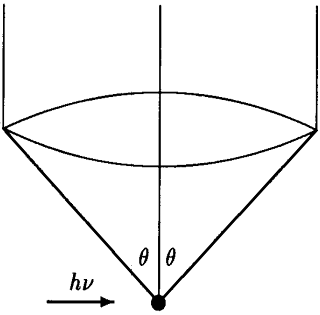

Figure 7.1: The Heisenberg microscope.

of the microscope is of finite width we do not know how much momentum was imparted. In fact the reflected photon can travel anywhere in the cone designated by the angle θ. If the wavelength of the photon is λ then the resolving power of the microscope is

$$\Delta x = \frac{\lambda}{\sin\theta} \ . \tag{7.7.44}$$

The momentum transferred to the electron (the uncertainty in the momentum) is given by:

$$\Delta p = p\sin\theta \tag{7.7.45}$$

where p is the momentum of the photon. But

$$p = \frac{h}{\lambda} \ . \tag{7.7.46}$$

Therefore,

$$\Delta x \Delta p \approx h \ . \tag{7.7.47}$$

Notice that this "derivation" implies that this inaccuracy is intrinsic and cannot be decreased by getting better instruments. We now turn to the mathematical proof. It is based on the following properties:

1) A and B are self-adjoint.

2) The expressions for ΔA and ΔB are:

$$(\Delta A)^2 = \langle (A - \langle A \rangle)^2 \rangle$$

$$(\Delta B)^2 = \langle (B - \langle B \rangle)^2 \rangle \ .$$

3) The Schwarz inequality,

$$|(f,g)|^2 \leq (f,f)(g,g) \ .$$

The statement to be proven is:

$$\Delta A \Delta B \geq |(\psi, [A, B]\psi)| \,. \tag{7.7.48}$$

We first define the operators

$$A' = A - \langle A \rangle \tag{7.7.49}$$

$$B' = B - \langle B \rangle \,. \tag{7.7.50}$$

These are still self-adjoint. Also,

$$[A', B'] = [A, B] \,. \tag{7.7.51}$$

All we have done is to subtract the mean values. Furthermore, we now have:

$$(\Delta A)^2 = (A'\psi, A'\psi) \tag{7.7.52}$$

since

$$\begin{aligned} (\Delta A)^2 &= (\psi, A'^2\psi) \\ &= (\psi, (A - \langle A \rangle)^2\psi) \\ &= \langle (A - \langle A \rangle)^2 \rangle \,. \end{aligned} \tag{7.7.53}$$

To obtain (7.7.48) we now just apply the Schwarz inequality after the following short computation

$$\begin{aligned} (\psi, [A, B]\psi) &= (\psi, [A', B']\psi) \\ &= (\psi, A'B'\psi) - (\psi, B'A'\psi) \\ &= (A'\psi, B'\psi) - (B'\psi, A'\psi) \\ &= (A'\psi, B'\psi) - (A'\psi, B'\psi)^* \\ &= 2i\Im(A'\psi, B'\psi) \,. \end{aligned} \tag{7.7.54}$$

Thus,

$$\begin{aligned} |(\psi, [A, B]\psi)| &= 2|\Im(A'\psi, B'\psi)| \\ &\leq 2|(A'\psi, B'\psi)| \end{aligned} \tag{7.7.55}$$

and by Schwarz's inequality

$$\begin{aligned} &\leq 2(A'\psi, A'\psi)^{1/2}(B'\psi, B'\psi)^{1/2} \\ &= 2\Delta A \Delta B \,. \end{aligned} \tag{7.7.56}$$

Thus, finally we find

$$\Delta A \Delta B \geq \frac{1}{2}|(\psi, [A, B]\psi)| \tag{7.7.57}$$

as stated. The proof just given is completely general except for the case where $[A, B] \neq 0$ and the state ψ is an eigenstate of either A or B , say

$$A\psi = a\psi \,. \tag{7.7.58}$$

The point here is that in this case, either ψ is not in the domain of B or else $B\psi$ is not in the domain of A. So, in either case an expression such as $(\psi, [A, B]\psi)$ is meaningless. To see this consider

$$\begin{aligned}(\psi, [A, B]\psi) &= (\psi, AB\psi) - (\psi, BA\psi) \\ &= (A\psi, B\psi) - a(\psi, B\psi) \\ &= a(\psi, B\psi) - a(\psi, B\psi) = 0 \end{aligned} \tag{7.7.59}$$

contrary to our assumption of a non-zero commutator.

In arriving at this contradiction we have also made heavy use of the fact that A is self-adjoint. For the case where ψ is indeed an eigenstate of say A one can either arrive at a general result by a different means [7.4] or else one can reconsider the whole question from the start by explicitly using that state.

We now proceed to discuss the time-evolution of a quantum state. In classical mechanics the evolution of a state is given by a set of first order differential equations, Hamilton's equations, and is therefore unique. The corresponding thing in quantum mechanics is Schrödinger's equation which is also first order in time and hence predicts the evolution of a state uniquely. Furthermore, the evolution of a state $\Psi(t, x)$ gives the time evolution of any observable's expectation value according to

$$\langle A \rangle_t = (\Psi(t, x), A\Psi(t, x)) \ . \tag{7.7.60}$$

On the other hand, the "value" of the observable does not evolve in a completely predictable fashion. Just as in classical mechanics the Hamiltonian plays a special role in determining the evolution so does the Hamiltonian operator in quantum mechanics. We now state this as our fifth and final assumption.

7.8 A5 - The Schrödinger Equation

For every physical system there exists an observable, the total energy, to which there corresponds a self-adjoint operator H called the Hamiltonian. The Hamiltonian determines the time evolution of the system according to the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi \tag{7.8.61}$$

provided the system is not disturbed.

Note, the last proviso states that if a measurement is made, (7.8.61) ceases to hold. This is due to the fact that making a measurement involves disturbing the system in an essentially unpredictable manner as we saw with the Heisenberg microscope. If the disturbance were predictable we could include it as some new, perhaps time-dependent, "potential" in the Hamiltonian and thus in principle predict the outcome of all experiments. This is not possible, however.

In order that our interpretation of quantum mechanics as given by A1, A2, A3 and A4 be consistent with A5 we need that if Ψ is normalized at some instant, that it should remain so. This implies that $(\Psi(t,x),\Psi(t,x))$ must be time-independent. But

$$\frac{d}{dt}\left(\Psi(t,x),\Psi(t,x)\right)=\left(\frac{\partial\Psi(t,x)}{\partial t},\Psi(t,x)\right)+\left(\Psi(t,x),\frac{\partial\Psi(t,x)}{\partial t}\right) \tag{7.8.62}$$

and using (7.8.61)

$$\begin{aligned}\frac{d}{dt}\left(\Psi(t,x),\Psi(t,x)\right) &= \left(\frac{-iH}{\hbar}\Psi(t,x),\Psi(t,x)\right)+\left(\Psi(t,x),\frac{-iH}{\hbar}\Psi(t,x)\right)\\ &= \frac{i}{\hbar}\left[(H\Psi(t,x),\Psi(t,x))-(\Psi(t,x),H\Psi(t,x))\right]\\ &= \frac{i}{\hbar}\left[(\Psi(t,x),H\Psi(t,x))-(\Psi(t,x),H\Psi(t,x))\right]\\ &= 0\end{aligned} \tag{7.8.63}$$

where we have used the fact that H is self-adjoint.

The way we have formulated the time-dependence via the Schrödinger equation allows us to consider the time t as a parameter labelling different states on the hilbert space $\mathcal{H}$. All that the Schrödinger equation does is to determine how one vector in $\mathcal{H}$ evolves into another. Thus, it is natural to look for an operator that performs this evolution. In fact, to preserve the inner product, the evolution operator must be unitary. Also if we compare (7.7.50) with our statement of Stone's Theorem we see that the Hamiltonian H is the generator of the evolution operator. Rather than use Stone's Theorem we shall derive this result directly. Thus, assume that there exists a unitary operator $U(t,t_0)$ such that

$$\Psi(t,x)=U(t,t_0)\Psi(t_0,x) \tag{7.8.64}$$

where

$$U(t,t_0)=1\,. \tag{7.8.65}$$

Substituting into (7.8.61) we get

$$\left[i\hbar\frac{\partial U(t,t_0)}{\partial t}-HU(t,t_0)\right]\Psi(t_0,x)=0\,. \tag{7.8.66}$$

This must be true for all possible initial states $\Psi(t_0,x)$, and hence

$$i\hbar\frac{\partial U(t,t_0)}{\partial t}=HU(t,t_0)\,. \tag{7.8.67}$$

The solution is

$$U(t,t_0)=\exp\left[-\frac{iH}{\hbar}(t-t_0)\right] \tag{7.8.68}$$

or

$$U(t,t_0)=1+\sum_{n=1}^{\infty}\frac{1}{n!}\left[\frac{iH(t-t_0)}{\hbar}\right]^n \tag{7.8.69}$$

where (7.8.69) gives a definition of the function exp. Notice that $U(t, t_0)$ is not an observable since it is not self-adjoint. Nevertheless it is a very physical object. We shall encounter this operator again when we examine the so-called transformation theory of quantum mechanics

7.9 Time Evolution: Constants of the Motion

The expectation value of an observable $\mathcal{A}$ is given by

$$\langle A\rangle_t = (\Psi(t,x), A\Psi(t,x)) \ . \tag{7.9.70}$$

Hence,

$$\begin{aligned} \frac{d}{dt}\langle A\rangle_t &= \frac{d}{dt}(\Psi(t,x), A\Psi(t,x)) \\ &= \left(\frac{\partial\Psi(t,x)}{\partial t}, A\Psi(t,x)\right) + \left(\Psi(t,x), A\frac{\partial\Psi(t,x)}{\partial t}\right) \ . \end{aligned} \tag{7.9.71}$$

Using (7.8.61) this becomes

$$\begin{aligned} \frac{d}{dt}\langle A\rangle_t &= \left[\left(-\frac{i}{\hbar}H\Psi(t,x), A\Psi(t,x)\right) + \left(\Psi(t,x), A(-\frac{i}{\hbar}H)\Psi(t,x)\right)\right] \\ &= \frac{i}{\hbar}(\Psi(t,x), (HA - AH)\Psi(t,x)) \end{aligned} \tag{7.9.72}$$

where we have used the self-adjointness of H. Thus,

$$\frac{d}{dt}\langle A\rangle_t = \frac{i}{\hbar}(\Psi(t,x), [H,A]\Psi(t,x)) \ . \tag{7.9.73}$$

This gives the evolution of the "mean value" of an observable and is sometimes very useful for evaluating $\langle A\rangle_t$. It is frequently simpler to solve this as a differential equation rather than to evaluate $\langle A\rangle_t$ directly. For an example see problem 7.5.

Now, suppose $[H, A] = 0$. Then (7.9.73) gives

$$\frac{d}{dt}\langle A\rangle_t = 0 \ . \tag{7.9.74}$$

Thus, the expectation value of $\mathcal{A}$ does not change in time and we can state that $\mathcal{A}$ is a constant of the motion. In a similar manner we can show that (Problem 7.8)

$$\frac{d}{dt}(\Delta A)_t = 0 \ . \tag{7.9.75}$$

An important special case of this is the Hamiltonian itself. Since $[H, H] = 0$ we immediately get that the Hamiltonian is a constant of the motion. But the Hamiltonian represents the total energy of the undisturbed system. Thus, we have shown that for an undisturbed system, the total energy does not change with time.

7.10 Time-Energy Uncertainty Relation

The inequality (7.6.40) states that for classically conjugate variables (i.e. those satisfying $[p,q] = i\hbar$) we have the Heisenberg uncertainty or indeterminacy relation

$$\Delta p \Delta q \geq \hbar/2 \,. \tag{7.10.76}$$

Such a relation holds also for the time and energy variables even though they do *not* satisfy a commutation relation

$$[E,t] = i\hbar \,. \tag{7.10.77}$$

This commutator might be conjectured from the fact that in the time-dependent Schrödinger equation the "energy operator" appears as

$$E = i\hbar\frac{\partial}{\partial t} \,. \tag{7.10.78}$$

As we now show equation (7.10.77) is *false* for any physical system. The reason for this is that the energy for a physical system must have a finite lower bound. Otherwise the system would wind up in the lowest state of negative energy at $E = -\infty$ (by radiating its energy away) and stay there forever. No perturbation would be strong enough to excite the system from $E = -\infty$ to a finite value. This being the case we must have a finite E_0 with all energies $E \geq E_0$ and thus there can be no self-adjoint extensions for an operator t satisfying the commutator equation (7.10.77). (See problem 7.10.) [1] Hence time could not be an observable if equation (7.10.77) were to hold.

We now show directly that if there exists a finite lower bound E_0 (the lowest energy eigenvalue of H) such that all eigenvalues E satisfy $E \geq E_0$ then we obtain a contradiction from equation (7.10.77) if both the Hamiltonian H (total energy) and the time t are to be observables.

Let ϕ_0 be the energy eigenfunction corresponding to the lowest energy E_0 .

$$H\phi_0 = E_0\phi_0 \,. \tag{7.10.79}$$

Now pick any positive frequency ω and define the operator

$$b = e^{i\omega t} \,. \tag{7.10.80}$$

Then, (see problem 7.12)

$$[H,b] = -\hbar\omega b \,. \tag{7.10.81}$$

So,

$$Hb\phi_0 = bH\phi_0 - \hbar\omega b\phi_0 \tag{7.10.82}$$

or

$$H(b\phi_0) = (E_0 - \hbar\omega)(b\phi_0) \,. \tag{7.10.83}$$

---

[1] If the energy is also bounded above, then t has a one-parameter family of self-adjoint extensions. In this case, however, the spectrum of the t-operator is discrete.

Therefore, either E_0 is not the lowest eigenvalue of H or else

$$b\phi_0 = 0\ .$$

But, if t is to be an observable then ωt must be self adjoint and $b = e^{i\omega t}$ must be unitary. As a consequence

$$b\phi_0 = 0 \quad \Rightarrow \quad \phi_0 = 0\ .$$

Thus, we have to conclude that the commutator (7.10.77) is false.

In spite of the absence of a "time-energy commutation relation", a time-energy uncertainty relation holds. This relation is extremely useful and it is important to understand what it means. It is for this reason we have given such a lengthy discussion above. We now proceed to the derivation of this relation.

The "time" involved in the time-energy uncertainty relation is not the time parameter in the Schrödinger equation, instead it is the "time of a process" associated with an observable $\mathcal{A}$. We therefore define the evolution time T_A associated with the observable $\mathcal{A}$ by

$$T_A = \frac{(\Delta A)_t}{\left|\frac{d\langle A\rangle_t}{dt}\right|}\ . \tag{7.10.84}$$

To see what this means consider the change in the expectation value $\Delta\langle A\rangle_t$ of the observable $\mathcal{A}$ in a time interval Δt. This is

$$\begin{aligned} |\Delta\langle A\rangle| &= |\langle A\rangle_{t+\Delta t} - \langle A\rangle_t| \\ &\approx \left|\frac{d\langle A\rangle}{dt}\right| \Delta t\ . \end{aligned} \tag{7.10.85}$$

Now for the change $|\Delta\langle A\rangle|$ to be measurable requires that $|\Delta\langle A\rangle|$ be at least as large as the uncertainty $(\Delta A)_t$ in the observable $\mathcal{A}$. This gives us the length of time Δt which we must wait to be able to "see" any change in $\Delta\langle A\rangle_t$. Equating $|\Delta\langle A\rangle|$ with $|\Delta\langle A\rangle_t|$ we find

$$\Delta t = \frac{(\Delta A)_t}{\left|\frac{d\langle A\rangle_t}{dt}\right|} = T_A\ . \tag{7.10.86}$$

But this is just the time T_A. So T_A is the time required for the expectation value of $\mathcal{A}$ to change by an amount equal to the uncertainty in $\mathcal{A}$.

Returning to the time-energy uncertainty relation we find according to (7.9.73) that

$$\begin{aligned} \left|\frac{d\langle A\rangle_t}{dt}\right| &= \frac{1}{\hbar}|(\Psi, [H, A]\Psi)| \\ &= \frac{1}{\hbar}|(\Psi, [H', A']\Psi)| \end{aligned} \tag{7.10.87}$$

where as before

$$\begin{aligned} H' &= H - \langle H\rangle \\ A' &= A - \langle A\rangle\ . \end{aligned} \tag{7.10.88}$$

Then, by the same sequence of steps used to obtain the general uncertainty relation (7.6.40), we get:

$$\left|\frac{d\langle A\rangle_t}{dt}\right| \leq \frac{2}{\hbar}\Delta H \Delta(A)_t \ . \tag{7.10.89}$$

Writing ΔE for ΔH and solving for T_A we find

$$\Delta E T_A \geq \hbar/2 \tag{7.10.90}$$

or writing Δt for T_A we have

$$\Delta E \Delta t \geq \hbar/2 \ . \tag{7.10.91}$$

This is the famous time-energy uncertainty relation. Its meaning is clear from the derivation. Namely, let ΔE be the uncertainty (RMS deviation from the mean) in the total energy and let Δt be the minimum time required for a measurable change to occur in a given observable evolving according to the Hamiltonian describing the total energy. Under these circumstances the relation

$$\Delta E \Delta t \geq \hbar/2 \tag{7.10.92}$$

holds.

7.11 Time Evolution of Probability Amplitudes

We now derive an equation that tells us how the probability amplitudes evolve in time. Since the Hamiltonian H is self-adjoint it has a complete set of eigenvectors $u_k(x)$. Assume these are normalized and therefore form a basis. Thus,

$$H\, u_k(x) = E_k\, u_k(x) \tag{7.11.93}$$

where E_k are the eigenvalues of H and thus the allowed values of the total energy of the system. A given eigenstate then evolves according to

$$\Psi_k(t,x) = e^{-iE_k t/\hbar}\, u_k(x) \ . \tag{7.11.94}$$

Now, suppose an operator A corresponding to an observable $\mathcal{A}$ has eigenvalues $\{a_n\}$ and eigenstates $\{\phi_n(x)\}$. If the state of the system is $\Psi(t,x)$ we can write the probability amplitude α_n for observing a_n as

$$\alpha_n(t) = (\phi_n(x), \Psi(t,x)) \ . \tag{7.11.95}$$

Thus, we get the expansion

$$\Psi(t,x) = \sum_n \alpha_n(t)\phi_n(x) \ . \tag{7.11.96}$$

Equation (7.11.94) is obviously a special case of this where the observable $\mathcal{A}$ is the total energy. From (7.11.95) we get

$$\begin{aligned}
\frac{d\alpha_n(t)}{dt} &= \left(\phi_n, \frac{\partial \Psi}{\partial t}\right) \\
&= -\frac{i}{\hbar}(\phi_n, H\Psi) \\
&= -\frac{i}{\hbar}\left(\phi_n, H\sum_m \alpha_m(t)\phi_m\right) \\
&= -\frac{i}{\hbar}\sum_m \alpha_m(t)\,(\phi_n, H\phi_m) \ . \qquad (7.11.97)
\end{aligned}$$

Thus,

$$\frac{d\alpha_n(t)}{dt} = -\frac{i}{\hbar}\sum_m \alpha_m(t)\,(\phi_n, H\phi_m) \ . \qquad (7.11.98)$$

For an arbitrary observable $\mathcal{A}$ this is as far as we can go. If, however, A is the Hamiltonian H so that $\phi_m = u_m$, then we get

$$\begin{aligned}
\frac{d\alpha_n(t)}{dt} &= -\frac{i}{\hbar}\sum_m \alpha_m(t)E_m\delta_{nm} \\
&= -\frac{i}{\hbar}E_n\alpha_n(t) \ . \qquad (7.11.99)
\end{aligned}$$

So

$$\alpha_n(t) = \alpha_n(0)\, e^{-iE_n t/\hbar} \ . \qquad (7.11.100)$$

This result follows of course immediately from (7.11.94) by recalling that

$$\alpha_n(t) = (\phi_n, \Psi) \ .$$

In fact, this is precisely how we previously introduced the concept of stationary state in connection with the Schrödinger equation.

We now use these results to examine the time development of observables that are constants of the motion. Thus, if $\mathcal{A}$ is a constant of the motion then

$$[A, H] = 0 \ . \qquad (7.11.101)$$

We have already found that this implies that $\langle A\rangle_t$ does not change in time and similarly $(\Delta A)_t$ is constant in time. Now since (7.11.101) is time independent and A and H are both self-adjoint with an assumed common dense domain, they have a common eigenbasis. Thus, both

$$H\,u_k = E_k\,u_k \qquad (7.11.102)$$

and

$$A\,\phi_n = a_n\,\phi_n \ . \qquad (7.11.103)$$

We can by relabelling (or reordering) arrange that

$$\phi_{n,k} = u_{k,n} \tag{7.11.104}$$

so that, as stated, the operators have a common eigenbasis. The probability of measuring a_n for a system in the state $\Psi(t,x)$ is as before given by

$$|\alpha_n(t)|^2 = |(\phi_{n,k},\Psi)|^2 = |(u_{k,n},\Psi)|^2 \,. \tag{7.11.105}$$

Thus

$$\begin{aligned}|\alpha_n(t)|^2 &= \left|\left(u_{k,n},\Psi(0,x)\,E^{-iE_nt/\hbar}\right)\right|^2\\ &= |(u_{k,n},\Psi(0,x))|^2\\ &= |(\phi_{n,k},\Psi(0,x))|^2\end{aligned} \tag{7.11.106}$$

and is independent of t. This means that all measurements we can perform on $\mathcal{A}$ do not change with time and so we are truly justified in calling $\mathcal{A}$ a constant of the motion whenever $[A,H]=0$.

We have now formulated non-relativistic quantum mechanics as a theory of linear operators on a Hilbert space. A predominant role is played by the Hamiltonian operator. Thus, a rule for obtaining this operator is desirable. Unfortunately aside from the considerations we gave in Chapter 3 not much more can be done in general. There are more formal ways of approaching this problem of making a classical observable into a self-adjoint operator on the Hilbert space of physical states. Basically the problem is as follows.

Find pairs of classically conjugate variables p_r , q_r. Then impose the formal algebraic relation

$$[p_r,q_s] = -i\hbar\,\delta_{rs} \,. \tag{7.11.107}$$

Now look for operators on a Hilbert space $\mathcal{H}$ such that these operators satisfy the relation (7.11.107) and are self-adjoint. Worded in a slightly different way this problem was considered by von Neumann who found the following result. Except for unitary transformations (which amount to a rotation of the Hilbert space and are thus of no physical consequence) the Schrödinger representation of the commutation relations (7.11.107) is unique. The Schrödinger representation is

$$p_r = \frac{\hbar}{i}\frac{\partial}{\partial x_r} \tag{7.11.108}$$

$$q_r = x_r \,. \tag{7.11.109}$$

Actually von Neumann showed that there is a slight generalization of (7.11.108) and (7.11.109) called a direct sum, but this is also of no importance to us. It essentially amounts to writing

$$p_r = \begin{pmatrix} \frac{\hbar}{i}\frac{\partial}{\partial x_r} & 0 \\ 0 & \frac{\hbar}{i}\frac{\partial}{\partial y_r}\end{pmatrix} \tag{7.11.110}$$

$$q_r = \begin{pmatrix} x_r & 0 \\ 0 & y_r \end{pmatrix} \tag{7.11.111}$$

or even larger diagonal matrices with different independent variables x, y, z etc. in each new row. We shall have no need for this. Now once we have the representation (7.11.108) and (7.11.109) we form operators as stated in Chapter 3, paying due attention to the possible ambiguities there considered.

7.12 Problems

7.1 A particle is in a state given at $t = 0$ by:

$$\psi = \frac{1}{3}u_0(x) + \frac{i\sqrt{2}}{3}u_1(x) - \frac{\sqrt{6}}{3}u_2(x)$$

where u_0, u_1, u_2 are simple harmonic oscillator eigenfunctions corresponding to the energies $1/2\hbar\omega$, $3/2\hbar\omega$ and $5/2\hbar\omega$; respectively.

a) What is the most likely value of the energy that will be found in a single observation on this system? What is the probability of finding this value?

b) What is the average of the energy that would be obtained if the experiment in part a) could be repeated many times? What is the probability of getting this value?

c) A measurement of the energy yields a value $3/2\hbar\omega$. The measurement is immediately repeated. What is the resultant value of the energy? What is the wave-function immediately after the second measurement?

d) What is the wavefunction of the undisturbed system after a time t has elapsed?

7.2 A free particle is located at $x = a$ at $t = 0$ i.e. its wavefunction at $t = 0$ is given by

$$\Psi(0, x) = \delta(x - a)\ .$$

Find the wavefunction for $t > 0$. This solution is called the free particle propagator.
Hint: To evaluate an integral of the form $\int_{-\infty}^{\infty} \exp\{i\lambda x^2 + i\beta x\}\, dx$ pretend that λ is $\lambda + i\epsilon$ with $\epsilon > 0$ so that the integral is convergent. Then complete the square in the exponent of the exponential and change variables. The $i\epsilon$ in your answer will allow you to decide whether to take the positive or negative square root. Finally let $\epsilon \to 0$.

7.3 Minimum Uncertainty wavefunction.
We have seen (7.6.43) that $\Delta x \Delta p \geq \hbar/2$. Assume $\langle x \rangle = \langle p \rangle = 0$. Now use the Schwarz inequality (6.2.4)

$$\| f \|^2 \| g \|^2 \geq |(f,g)|^2$$

and put

$$f = x\psi \quad , \quad g = p\psi \ .$$

Show that the equality in the uncertainty will hold only if

$$p = \lambda \, x\psi$$

with λ a constant and

$$(\psi, (xp + px)\psi) = 0 \ .$$

Hence derive an equation for ψ and solve it explicitly.

7.4 A free particle of mass m is at $t = 0$ in a state described by

$$\Psi(0,x) = \left[2\pi L^2\right]^{-1/4} \exp -(x/2L)^2 \ .$$

What is the wavefunction for an arbitrary time $t > 0$? Compute the uncertainties Δx and Δp as functions of time. This illustrates the "spreading" of a wave packet.
Hint: Look at problem 7.2.

7.5 A free particle is, at $t = 0$, in a state described by the wavefunction

$$\Psi(0,x) = \begin{cases} A \sin^2 \frac{\pi x}{a} & |x| < a \\ 0 & |x| > a \end{cases} \ .$$

Find for $t > 0$ the following expectation values

$$\langle p \rangle_t \ , \ \langle x \rangle_t \ , \ \langle p^2 \rangle_t \ , \ \langle \frac{d}{dt}(xp + px) \rangle_t \ , \text{ and finally } \langle x^2 \rangle_t$$

Hint: Use equation (7.9.73). If you try to evaluate these results using the time-dependent wavefunction $\Psi(t,x)$ you will get some impossible integrals.

7.6 Consider a particle under the influence of an Hamiltonian

$$H = \frac{p^2}{2m} + V(x)$$

so that

$$i\hbar\frac{\partial\Psi}{\partial t} = H\Psi\,.$$

Show that if $\langle x\rangle$ is the "centre of mass" of the wave packet and $\langle p\rangle$ the average momentum of the particle then

$$\frac{d\langle x\rangle}{dt} = \frac{\langle p\rangle}{m}$$

and

$$\frac{d\langle p\rangle}{dt} = \langle F(x)\rangle = -\langle\frac{dV(x)}{dx}\rangle\,.$$

These are known as the *Ehrenfest Equations.* To be equivalent to Newton's equations requires that

$$\langle F(x)\rangle = F(\langle x\rangle)\,.$$

Discuss under what circumstances this condition is approximately valid.

7.7 Prove the compatibility theorem.

7.8 Show that if $[H, A] = 0$ then ΔA does not change in time.

7.9 Consider a free electron described at time $t = 0$ by the Gaussian wavepacket

$$\Psi(0,x) = \left[2\pi L^2\right]^{-1/4}\exp -(x/2L)^2\,.$$

Using the experience from problem 7.4 answer the following questions:

a) What is $\langle x\rangle$ at any time?

b) What is Δx as a function of time?

This is known as the spreading of a wave packet. If this packet corresponds to an electron with 20 keV energy and a width of 100 Å at $t = 0$, what is its width after travelling 100 m?

7.10 Consider the space of functions of E belonging to $\mathcal{L}_2(0,\infty)$. The relationship

$$[E, t] = i\hbar$$

can be represented on this space by

$$E\,f(E) = E\,f(E)$$

and

$$t\,f(E) = -i\hbar\,\frac{df(E)}{dE}\ .$$

Show that the operator t so defined can not be an observable, i.e. that it has no self-adjoint extensions. This proves that if the energy has a lower bound, a relationship such as (7.10.77) cannot hold if time is to be an observable.

7.11 A particle is in a state described by the unnormalized wave function:

$$f(x) = Ae^{-a|x|} \quad a > 0\ .$$

Find the length of an interval around the origin such that the probability of finding the particle in this interval is 40 % .

7.12 Show that for any two operators A, B such that

$$[A, B] = c \quad \text{a c-number}$$

then

$$[A, e^{B}] = c\,e^{B}\ .$$

Hint: Expand e^B and use the fact that

$$[A, B^n] = ncB^{n-1}\ .$$

Bibliography

[7.1] D.T. Gillespie, *A Quantum Mechanics Primer* - International Textbook Co. Scranton, Pa. (1970). This is a very readable little book describing the orthodox interpretation of quantum mechanics.

[7.2] One of the best books is still: P.A.M. Dirac, *The Principles of Quantum Mechanics* - Oxford at the Clarendon Press, 4th edition (1958).

[7.3] A book containing numerous physical examples and written by one of the founders of the subject is: W. Heisenberg, *Physical Principles of the Quantum Theory* - Dover Publishing, Inc., New York.

[7.4] P. Carruthers and M. M. Nieto, Rev. Mod. Phys. **40**, 411- 440, (1968).

Chapter 8

Distributions and Fourier Transforms

8.1 Introduction

In this chapter we develop some more mathematical tools. Again these tools are not important for computational purposes, but they are important as a justification for the calculations performed in practice. Although physicists need not, as a rule, practice mathematics with the same rigour as a mathematicians they need to know mathematics sufficiently well to know when it is safe to be sloppy. Furthermore in areas such as quantum field theory where it is not known whether the difficulties encountered are due to bad mathematics, or bad physics, or both it is important to ensure that the mathematics at least, is correct. To this end, we give here a brief introduction to some of the results of modern analysis. The presentation, although still at a submathematical level is intended for the more mathematically inclined student. We give definitions and theorems, but the proofs for the theorems are only sketched, or omitted altogether. To compensate for this we list several relevant references at the end of the chapter.

8.2 Functionals

A function may be considered as a mapping from a certain well-defined set of numbers called the domain into another set of numbers called the range. Thus, if f denotes a function then $f(x)$ denotes the value of the function f at the point x. This distinction is not always made but clearly there is such a distinction. We shall now consider a mapping whose domain is a set of functions called test functions and whose range is the set of real numbers. Such a mapping is called a functional. If T is a functional then $T(f)$ is the *value* of the functional T at the function f. Thus the arguments of functionals are functions. From the class of all possible functionals we pick out a particularly simple class, namely the linear functionals. A functional T is linear if for f and g belonging to the domain of T and a, b two numbers

$$T(af + bg) = aT(f) + bT(g) . \tag{8.2.1}$$

An example of such a functional is

$$T(f) = \int_{-\infty}^{\infty} t(x)\, f(x)\, dx \tag{8.2.2}$$

where $t(x)$ is a fixed function and $f(x)$ is in the domain of T if the right hand side is convergent. Furthermore a functional T is bounded if for all f in a given space $|T(f)| \leq c \parallel f \parallel$ where c is a positive constant. There is a remarkable theorem for bounded linear functionals on a Hilbert space.

Riesz Representation Theorem

Let $\mathcal{H}$ be a Hilbert space and T a bounded linear functional on $\mathcal{H}$. Then there exists a uniquely determined vector f_T of $\mathcal{H}$ such that

$$T(g) = (f_T, g) \tag{8.2.3}$$

for all $g \in \mathcal{H}$. Conversely, of course, any vector $f \in \mathcal{H}$ defines a bounded linear functional T_f by

$$T_f(g) = (f, g) . \tag{8.2.4}$$

Proof

The proof is rather straightforward and is a proof by construction. Uniqueness is obvious. For suppose f' is another vector besides f_T satisfying (8.2.3), then

$$(f' - f_T, g) = 0 \tag{8.2.5}$$

for all $g \in \mathcal{H}$. Thus $f' - f_T = 0$ as desired. To prove that f_T exists consider the null space N_T of T where

$$N_T = \{g \in \mathcal{H} \mid T(g) = 0\} . \tag{8.2.6}$$

If $N_T = \mathcal{H}$ take $f_T = 0$. This is the trivial case. Now assume $N_T \neq \mathcal{H}$. Then there exists at least one vector $f_0 \neq 0$ belonging to $N_T^{\perp}$, the orthogonal complement of N_T . In this case define

$$f_T = \frac{T(f_0)^*}{\parallel f_0 \parallel^2}\, f_0 . \tag{8.2.7}$$

This is the desired f_T as we now prove. Suppose $g \in N_T$. Then

$$T(g) = 0 = (f_T, g) . \tag{8.2.8}$$

Next, if g is of the form

$$g = \alpha f_0 \tag{8.2.9}$$

then we have

$$(f_T, g) = (f_T, \alpha f_0) = \alpha T(f_0) = T(g) \tag{8.2.10}$$

as required. We now show that any $g \in \mathcal{H}$ can be written

$$g = \alpha f_0 + \beta f_1 \tag{8.2.11}$$

where $f_1 \in N_T$. To prove this recall that

$$T(f_T) \neq 0 \, . \tag{8.2.12}$$

Then we have the identity

$$g = \left(g - \frac{T(g)}{T(f_T)} f_T \right) + \frac{T(g)}{T(f_T)} f_T \tag{8.2.13}$$

which is of the form (8.2.11). Thus, since T is linear we have completed our proof and shown that

$$T(g) = (f_T, g) \tag{8.2.14}$$

for all $g \in \mathcal{H}$. This shows that on a Hilbert space the only linear functionals are those given by inner products. We want to extend this notion somewhat. Therefore, it is natural that we must go beyond the concept of Hilbert space.

In general to define a space we must have a criterion for deciding when two points of the space are "close". This criterion defines the *topology* of the space. For example, in the finite dimensional vector spaces $\mathcal{E}_n$ we use the Euclidean norm $(x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$ to measure closeness. In Hilbert space we use the norm

$$\| f \| = (f, f)^{1/2} \tag{8.2.15}$$

to measure closeness. For functions one also frequently uses point-wise estimates of the form $|f(x) - g(x)|$. All of these criteria are useful and define different topologies. For functionals one also has an estimate which is derived by analogy with (8.2.2). Thus, if T and S are bounded linear functionals, meaning that there are positive constants c, c' such that

$$T(f) \leq c \, \| f \| \tag{8.2.16}$$

and

$$S(f) \leq c' \, \| f \| \tag{8.2.17}$$

for all f in a given space $\mathcal{X}$, then T and S are "close" if $|T(f) - S(f)|$ is small. Here $\| f \|$ denotes the appropriate norm in $\mathcal{X}$. Thus, the notion of "close" (or topology) of the linear functionals on $\mathcal{X}$ is derived from the topology of $\mathcal{X}$ itself.

Dual Space

Let $\mathcal{X}$ be a space of functions with a given topology. Now consider the set $\mathcal{X}'$ of all bounded linear functionals on $\mathcal{X}$. Then $\mathcal{X}'$ is itself a linear vector space with the topology of $\mathcal{X}'$ determined by the topology of $\mathcal{X}$. We call $\mathcal{X}'$ the dual space of $\mathcal{X}$.

An example of these concepts is Hilbert space itself. In this case the dual of $\mathcal{H}$ is $\mathcal{H}$ itself. In fact it is logically correct to consider the inner product on a Hilbert space as being formed by elements from two spaces, the Hilbert space $\mathcal{H}$ and its dual, which is of course a copy of $\mathcal{H}$.

The point of all this is that one can take linear functionals that are as singular or pathological as one wishes if it is possible to find a space of functions sufficiently nice to compensate for these pathologies. The space of nice functions is called the space of test-functions. There are many test-function spaces. One of the most useful of these is the Schwarz space $\mathcal{S}$. Its dual space is called $\mathcal{S}'$, the space of *tempered distributions* and is sufficiently general to encompass almost any kind of "function" we shall encounter. To describe these spaces we need some more terminology.

A function with continuous derivatives up to and including the nth is called C^n. Thus, continuous functions are called C^0. If a function is C^n for all n it is called C^∞. Using this terminology we can define $\mathcal{S}$ as the space of all C^∞ functions which together with their derivatives vanish at infinity faster than the inverse of any polynomial. To make this more explicit we define the sequence of semi-norms [1]

$$\| f \|_{r,n} = \sup_x \left| x^r \frac{d^n f}{dx^n} \right| \tag{8.2.18}$$

where "sup" means "least upper bound". In that case f belongs to $\mathcal{S}$ iff

$$\| f \|_{r,n} < \infty \tag{8.2.19}$$

for all integers r, n. This specifies the topology or notion of closeness in $\mathcal{S}$ So, for example, a sequence $\{f_j\}$ of functions in $\mathcal{S}$ converges to f if for each r and n

$$\lim_{j\to\infty} \| f_j - f \|_{r,n} = 0\,. \tag{8.2.20}$$

In terms of this the *tempered distributions* also have a topology whose definition can be made very similar to the ϵ, δ definition for ordinary functions. Thus, T is continuous at f_0 if given an $\epsilon > 0$ there exist integers r, n and a $\delta > 0$ such that for

$$\| f - f_0 \|_{r,n} < \delta \tag{8.2.21}$$

we have

$$|T(f) - T(f_0)| < \epsilon\,. \tag{8.2.22}$$

One way to ensure that (8.2.22) follows from (8.2.21) is to insist that for all r, n there exists a positive constant c such that

$$|T(f)| \le c \,\| f \|_{r,n} \tag{8.2.23}$$

---

[1] The difference between a semi-norm and a norm is that a semi-norm may vanish for a given element even though that element is different from zero.

since then

$$|T(f) - T(f_0)| = |T(f - f_0)| \leq c \parallel f - f_0 \parallel_{r,n} \quad . \tag{8.2.24}$$

In fact, there is a theorem that states that every *continuous linear functional* T on $\mathcal{S}$ satisfies (8.2.23). This means that one can use (8.2.23) to define the topology on $\mathcal{S}'$.

To prepare us for future applications we introduce two more notations for distributions. To specify the value of T at f we have used $T(f)$. We can also write this (T, f). This does not mean we have an inner product, it is simply another way of writing $T(f)$. As a matter of fact as physicists we carry this even a step further and write this as

$$T(f) \equiv \int_{-\infty}^{\infty} T(x)\, f(x)\, dx \quad . \tag{8.2.25}$$

Again this is a purely symbolic way of writing $T(f)$ and does not imply that any integral such as (8.2.25) exists in any of the usual senses of integral. Nevertheless, the notation (8.2.25) is extremely suggestive and thus if applied with due caution one may treat this expression as an integral.

The most common of the distributions so treated is the δ function. It is defined by

$$\delta(f) = f(0) \quad . \tag{8.2.26}$$

On the other hand we frequently write this as

$$\int_{-\infty}^{\infty} \delta(x)\, f(x)\, dx = f(0) \, . \tag{8.2.27}$$

It is an easy matter to prove that no function δ with the property (8.2.27) can exist. [2] However, if we realize that (8.2.27) does not imply a genuine integral of a function δ and is just another, but very suggestive, way of writing (8.2.26) then all objections to writing (8.2.27) are removed. That δ is not a function can also be seen from the fact that although functions may be multiplied by functions to give functions, distributions cannot generally be multiplied by distributions or functions. For example, if we consider the product of $1/x$ and $\delta(x)$ this is not defined in general. Nevertheless there is a smaller domain for which this product makes sense. An even more acute example is the product

$$\delta(x)\delta'(x) \tag{8.2.28}$$

where $\delta'(x)$ is the derivative of $\delta(x)$.

We now define differentiation of distributions. In fact the definition is given by analogy with integration by parts using (8.2.25). Thus, we define

$$\frac{d^n T}{dx^n}(f) = (-1)^n\, T\left(\frac{d^n f}{dx^n}\right) \quad . \tag{8.2.29}$$

---

[2] See von Neumann's book [8.4], pages 23-25.

This expression is obviously well-defined for all $T \in \mathcal{S}'$ since if $f \in \mathcal{S}$ so is $\frac{d^n f}{dx^n} \in \mathcal{S}$. In the notation (8.2.25) the definition of the derivative reads

$$\int_{-\infty}^{\infty} \frac{d^n T}{dx^n} f(x)\, dx = (-1)^n \int_{-\infty}^{\infty} T(x) \frac{d^n f}{dx^n}\, dx \,. \tag{8.2.30}$$

It is a simple matter to generalize these results to test functions of several variables and distributions over these variables. Thus, if $f(x_1, x_2, \ldots, x_k)$ is an element of $\mathcal{S}^{(k)}$ in each variable, then we may have a distribution T in the dual space $\mathcal{S}'^{(k)}$ such that $T(f)$ is well defined. Again another possible symbolic notation for $T(f)$ would be

$$T(f) = \int T(x_1, x_2, \ldots, x_k)\, f(x_1, x_2, \ldots, x_k)\, dx_1 dx_2 \ldots dx_k \quad . \tag{8.2.31}$$

We emphasize once more that although (8.2.31) looks like an integral it is *not*. This is simply a symbolic way of writing $T(f)$. Nevertheless we shall use this way of writing almost all the time since it is the standard notation for physicists.

8.3 Fourier Transforms

Consider the linear transformations $\mathcal{F}$ and $\bar{\mathcal{F}}$ defined on $\mathcal{S}$ according to

$$(\mathcal{F}f)(p) = \int_{-\infty}^{\infty} e^{-ipx} f(x)\, dx \equiv F(p) \tag{8.3.32}$$

$$(\bar{\mathcal{F}}F)(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ipx} F(p)\, dp \,. \tag{8.3.33}$$

Clearly (8.3.32) defines a uniformly and absolutely convergent integral since e^{-ipx} can only improve the convergence of an already splendidly convergent integral. We shall now prove that $\mathcal{F}$ and $\bar{\mathcal{F}}$ map $\mathcal{S}$ onto $\mathcal{S}$ in a continuous one to one manner. The proof will give us as a side benefit the formal result

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ip(x-y)}\, dp = \delta(x-y) \,. \tag{8.3.34}$$

Consider the expression

$$\lim_{\epsilon \to 0+} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ipx}\, dp\, e^{-\epsilon p^2} \int_{-\infty}^{\infty} e^{-ipy} f(y)\, dy \,. \tag{8.3.35}$$

Now, for $\epsilon > 0$ both integrals exist and we may interchange their order. Furthermore,

$$\int_{-\infty}^{\infty} e^{ipx - \epsilon p^2}\, dp = \sqrt{\frac{\pi}{\epsilon}} \exp\{-x^2/4\epsilon\} \,. \tag{8.3.36}$$

Thus, we get for (8.3.35)

$$\lim_{\epsilon \to 0+} \frac{1}{\sqrt{4\pi\epsilon}} \int_{-\infty}^{\infty} e^{-(x-y)^2/4\epsilon} f(y)\, dy \,. \tag{8.3.37}$$

Now consider a circle $(x-y)^2 = R^2$. Clearly, due to the factor $e^{-(x-y)^2/4\epsilon}$ any contribution to the integral (8.3.37) from points outside the circle vanishes in the limit as $\epsilon \to 0$. Thus, we can estimate the difference between $f(x)$ and (8.3.37) by

$$\left| \frac{1}{\sqrt{4\pi\epsilon}} \int_{|x-y \leq R} e^{-(x-y)^2/4\epsilon} [f(y) - f(x)]\, dy \right| \\ \leq \sup_{|x-y| \leq R} |f(x) - f(y)| \to 0 \text{ as } R \to 0 \ . \tag{8.3.38}$$

This justifies (8.3.34) and shows that

$$\bar{\mathcal{F}}\mathcal{F} = 1 \ . \tag{8.3.39}$$

Using (8.3.34) we now also get

$$\mathcal{F}\bar{\mathcal{F}} = 1 \ . \tag{8.3.40}$$

Also we have that $\bar{\mathcal{F}}$, $\mathcal{F}$ map $\mathcal{S}$ onto $\mathcal{S}$ as stated.

Now, in mapping $\mathcal{S}$ onto a copy of itself using $\mathcal{F}$ what happens to $\mathcal{S}'$? In order to keep things well-defined, $\mathcal{S}'$ must be mapped onto $\mathcal{S}'$. Using the symbolic notation of (8.2.25) this is trivial to see. Since $\mathcal{F}$ is a unitary operator on Hilbert space we have for $f, g \in \mathcal{H}$ that

$$(\mathcal{F}f, \mathcal{F}g) = (f, g) \ . \tag{8.3.41}$$

This is known as *Parseval's theorem* and written out reads

$$\int_{-\infty}^{\infty} F^*(p)\, G(p)\, dp = \int_{-\infty}^{\infty} f^*(x)\, g(x)\, dx \tag{8.3.42}$$

or

$$\int_{-\infty}^{\infty} F^*(p)\, dp \int_{-\infty}^{\infty} e^{-ipx}\, g(x)\, dx = \int_{-\infty}^{\infty} g(x)\, dx \int_{-\infty}^{\infty} e^{-ipx}\, F^*(p)\, dp \tag{8.3.43}$$

where we have used the formulae defining F and G in terms of f and g and vice-versa. (8.3.43) is already in the desired form to define the Fourier transform of distributions. Thus suppose $g \in \mathcal{S}'$ then (8.3.43) reads

$$(\mathcal{F}g)\,(F^*) = g\left((\mathcal{F}F)^*\right) \ . \tag{8.3.44}$$

Thus we *define* the Fourier transform of distributions in $\mathcal{S}'$ using (8.3.44). In other words if $T \in \mathcal{S}'$ then the Fourier transform $\mathcal{F}T$ is defined by

$$(\mathcal{F}T)\,(f) = T\left((\mathcal{F}f)\right) \tag{8.3.45}$$

where $T \in \mathcal{S}$.

It is now a simple matter to use (8.3.45) to show that the Fourier transform maps $\mathcal{S}'$ onto $\mathcal{S}'$.

8.4 Rigged Hilbert Spaces

To motivate the use and definition of *rigged Hilbert spaces* (also known as *Gel'fand Triples*), we begin by considering the following eigenvalue problem on $\mathcal{H} = \mathcal{L}_2(-\infty, \infty)$

$$p\, u_k = \hbar k\, u_k \; . \tag{8.4.46}$$

Since

$$p = \frac{\hbar}{i}\frac{d}{dx} \tag{8.4.47}$$

we get

$$u_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx} \; . \tag{8.4.48}$$

Now the operator p corresponds to a physical observable, the momentum and hence the eigenvalue problem (8.4.46) has a definite physical meaning. It tells us what the possible results of measurements of p are and is also supposed to give the probability amplitude for obtaining a given measurement. Nevertheless the "eigenfunctions" $u_k(x)$ are not square-integrable and hence do not belong to our Hilbert space. This is an undesirable situation. It can of course be obviated by forming wave packets. However the plane waves (8.4.48) are particularly convenient for practical calculations and we would be reluctant to have to give up using them. Thus, we are tempted to enlarge our state vector space beyond Hilbert space. Actually this also provides many simplifications in the analysis of operators. However we shall not study that aspect.

To show one possible extension we first note that the functions $u_k(x)$ belong to $\mathcal{S}'$ if we define them as distributions in the following manner

$$u_k(f) = \int_{-\infty}^{\infty} u_k(x)\, f(x)\, dx \; . \tag{8.4.49}$$

Why did we choose $\mathcal{S}'$? The reasons are mainly technical. Thus $\mathcal{F}\mathcal{S}'$ is again $\mathcal{S}'$ and this is desirable. Actually other spaces of distributions may be used, but for the sake of concreteness we concentrate only on $\mathcal{S}'$. Now how does considering u_k as an element of $\mathcal{S}'$ help? To answer this we start with a definition.

Let A be a linear operator in $\mathcal{S}$. This means that A is also a linear operator in $\mathcal{H}$. In fact if A has a dense domain in $\mathcal{S}$ it has a dense domain in $\mathcal{H}$ since $\mathcal{S}$ is dense in $\mathcal{H}$. To see this consider the hermite functions $H_n(x)e^{-x^2/2}$. All of these are in $\mathcal{S}$ and any element in $\mathcal{L}_2(-\infty, \infty)$ can be approximated by linear combinations of these functions. Thus $\mathcal{S}$ is dense in $\mathcal{H}$.

Now given such an operator A then $T \in \mathcal{S}'$ is called a *generalized eigenvector* of A corresponding to the eigenvalue λ if

$$T(Af) = \lambda T(f) \tag{8.4.50}$$

for all $f \in \mathcal{S}$. Notice that by definition $u_k(x)$ is a generalized eigenvector of the momentum operator $p = \hbar/i\, d/dx$ since in the notation (8.2.25) we have

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{ikx} \frac{\hbar}{i} \frac{df}{dx}\, dx = -\hbar k \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{ikx} f(x)\, dx \tag{8.4.51}$$

for $f \in \mathcal{S}$. We have simply "integrated" by parts. Thus, we can now legitimately consider functions such as $u_k(x)$ as generalized eigenvectors. We still have to tie this together with the concept of Hilbert space. One more example is in order first.

Formally, the eigenvalue problem

$$x\, g_a(x) = a\, g_a(x) \tag{8.4.52}$$

has the solution

$$g_a(x) = \delta(x - a) \ . \tag{8.4.53}$$

Clearly $\delta(x-a)$ is not square integrable and hence is not in our Hilbert space $\mathcal{H}$. But for $f \in \mathcal{S}$ we have

$$\int_{-\infty}^{\infty} x\, \delta(x-a)\, f(x)\, dx = a \int_{-\infty}^{\infty} \delta(x-a)\, f(x)\, dx \ . \tag{8.4.54}$$

Thus $\delta(x-a)$ is a generalized eigenvector of the position operator x. We now define our *rigged Hilbert space.*

We begin with the space $\mathcal{S}$. On $\mathcal{S}$ are defined a countable sequence of norms $\| f \|_{n.r}$. We now also define on $\mathcal{S}$ an inner product which coincides with the $\mathcal{L}_2$ inner product. Now as stated $\mathcal{S}$ is dense in $\mathcal{L}_2$ and we identify $\mathcal{H}$ with $\mathcal{L}_2$. Thus $\mathcal{S}$ is identified as a subset of $\mathcal{H}$. Together with $\mathcal{S}$ and $\mathcal{H}$ we consider the space $\mathcal{S}'$. The triplet of spaces

$$\mathcal{S} \ , \ \mathcal{H} \ , \ \mathcal{S}'$$

form a rigged Hilbert space. It is usually denoted by

$$\mathcal{S} \subset \mathcal{H} \subset \mathcal{S}' \ .$$

The advantage of the symbolic notation (8.2.11) is now obvious. Thus, "inner products" exist between elements of $\mathcal{H}$ and $\mathcal{H}$ and elements of $\mathcal{S}$ and $\mathcal{S}'$. We do not form inner products between elements of $\mathcal{S}'$ and $\mathcal{S}'$. This is all about rigged Hilbert spaces that we shall need. It is sufficient to provide a justification of most of the manipulations that we shall carry out. Further details are readily available in the references. From now on we shall proceed as if "functions" like $\delta(x)$ and e^{ikx} were elements of $\mathcal{H}$. To justify our manipulations we can always fall back on the concept of rigged Hilbert spaces, but we shall not explicitly do so. As stated at the beginning, this chapter was simply to show that our formal manipulations can be fully justified.

8.5 Problems

8.1 Show that the appropriate normalization for the positive parity solution for scattering from a square well equation (5.7.94) to yield δ-function normalization is $1/\sqrt{\pi}$. You will have to use the continuity of the wavefunction at $x = \pm a$ as well as equation (5.7.101).

8.2 Show that T is a tempered distribution if T is defined by

$$T(f) = \sum_{k=0}^{m} \int_{-\infty}^{\infty} F_k(x) \frac{d^k f(x)}{dx^k} \, dx$$

where F_k are continuous functions bounded by

$$|F_k(x)| \leq C_k(1 + |x|^j)$$

for some C_k and j depending on k. As a matter of fact every tempered distribution can be written in this form. Symbolically one then writes

$$T = \sum_{k=0}^{m} (-1)^k \frac{d^k F_k(x)}{dx^k} \, .$$

This formula cannot be taken literally however since the $F_k(x)$ need not be differentiable. It arises from a formal integration by parts of the first equation above.

8.3 The test function space $\mathcal{D}$ consists of the space of $C^{(\infty)}$ functions of bounded support. The support of a function f, (supp f) is the complement of the largest open set on which the function vanishes. Show that if $\tilde{f} \in \mathcal{FD}$ then $\tilde{f}$ is an entire function.

8.4 Prove the Theorem: The Fourier transform of a *tempered distribution of fast decrease* is a $C^{(\infty)}$ function bounded by a polynomial. A tempered distribution of fast decrease F is of the form

$$F = fT$$

where $f \in \mathcal{S}$ and T is also a tempered distribution.
Hint: To prove that the Fourier transform of F is bounded by a polynomial use the result of problem 8.2.

8.5 Let $f(z)$ be an entire function vanishing rapidly at large $|\Re(z)|$. Show that

$$\lim_{\epsilon \to 0+} \frac{1}{2} \int_{-\infty}^{\infty} \left[\frac{1}{x - a + i\epsilon} + \frac{1}{x - a - i\epsilon} \right] f(x) \, dx = P \int_{-\infty}^{\infty} \frac{f(x)}{x - a} \, dx$$

where the principle value integral is defined by

$$P\int_{-\infty}^{\infty}\frac{f(x)}{x-a}\,dx=\lim_{\epsilon\to 0+}\left[\int_{-\infty}^{a-\epsilon}\frac{f(x)}{x-a}\,dx+\int_{a+\epsilon}^{\infty}\frac{f(x)}{x-a}\,dx\right]\ .$$

Furthermore, show that

$$\lim_{\epsilon\to 0+}\frac{\epsilon/\pi}{x^2+\epsilon^2}=\delta(x)\ .$$

Hence conclude that considered as distributions

$$\lim_{\epsilon\to 0+}\frac{1}{x-a\pm i\epsilon}=P\,\frac{1}{x-a}\mp i\pi\delta(x-a)$$

that is,

$$\lim_{\epsilon\to 0+}\int_{-\infty}^{\infty}\frac{f(x)\,dx}{x-a\pm i\epsilon}=P\int_{-\infty}^{\infty}\frac{f(x)\,dx}{x-a}\mp i\pi f(a)\ .$$

8.6 Using the result of problem 8.4 and defining

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{ikx}\,dk=\lim_{\epsilon\to 0+}\frac{1}{2\pi}\left[\int_{0}^{\infty}e^{ik(x+i\epsilon)}\,dk+\int_{-\infty}^{0}e^{ik(x-i\epsilon)}\,dk\right]$$

prove that

$$\int_{-\infty}^{\infty}e^{ikx}\,dk=\delta(x)\ .$$

8.7 Let $f(k)$ be a $C^{(\infty)}$ function bounded by a polynomial. Show that

$$F(z)=\int_{-\infty}^{\infty}f(k)\,e^{ikx}\,dk$$

is an entire function for $\Im(z)>0$. Using this and the result of the Theorem proved in problem 8.3 show that every tempered distribution is the boundary value of an analytic function.

8.8 Calculate the Fourier transform of $\delta^{(n)}(x)$.

8.9 Show that

$$x^m\delta^n(x)=\begin{cases}0 & \text{if}\quad n<m\\ (-1)^m\,m!\,\delta(x) & \text{if}\quad n=m\\ (-1)^m\frac{n!}{(n-m)!}\,\delta^{(n-m)}(x) & \text{if}\quad n>m\end{cases}\ .$$

Bibliography

[8.1] The standard reference on Rigged Hilbert Spaces is: I.M. Gel'fand, M.I. Graev, and N.Ya. Vilenkin, *Generalized Functions Vol.4*, Academic Press, N.Y. (1964), Chapters 1.3 and 1.4

[8.2] There are numerous other books on distribution theory of which we list several below. I.M. Gel'fand and G. E. Shilov, *Generalized Functions Vol.1* - Academic Press, N.Y. (1964).
A.H. Zemanian, *Distribution Theory and Transform Analysis* - McGraw -Hill Book Co. (1965)
H. Bremmerman, *Distributions, Complex Variables, and Fourier Transforms* - Addison-Wesley Publishing Co., Inc., Reading Mass. (1965).

[8.3] The original reference for the subject is: L. Schwarz, *Théorie des Distributions*, Hermann, Paris, Part I (1957), Part II (1959).

[8.4] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* - Princeton University Press (1955).

Chapter 9

Algebraic Methods

9.1 Introduction

Having developed all the tools we need we now turn to "practical" applications of these concepts. The first problem we reconsider is the simple harmonic oscillator.

We then move on to consider the rigid rotator in one and three dimensions. This allows us to introduce the concept of angular momentum. Next we solve the angular momentum eigenvalue problem analytically as well as algebraically. After that we discuss rotational invariance.

The algebraic solution of the angular momentum eigenvalue problem yields half odd-integral as well as integral eigenvalues. These half odd-integral eigenvalues correspond to a new degree of freedom of a particle, called spin. This is the final topic of discussion in this chapter.

9.2 Simple Harmonic Oscillator

The Hamiltonian for the simple harmonic oscillator is

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 \ . \tag{9.2.1}$$

As before, we introduce the classical angular frequency

$$\omega = \sqrt{k/m} \, . \tag{9.2.2}$$

Furthermore, as always we have

$$[x, p] = i\hbar \, . \tag{9.2.3}$$

We now define dimensionless operators

$$P = \frac{1}{\sqrt{m\hbar\omega}}p \tag{9.2.4}$$

$$Q = \sqrt{\frac{m\omega}{\hbar}}x \tag{9.2.5}$$

and get

$$H = \frac{1}{2}\hbar\omega\left(P^2 + Q^2\right) . \tag{9.2.6}$$

The commutator between P and Q is easily computed and yields

$$[P, Q] = -i . \tag{9.2.7}$$

So far we haven't done anything except some algebra to introduce dimensionless operators. Unlike earlier when we solved this problem by solving a differential equation, we now solve it using operator algebra techniques. To this end we introduce two non-hermitian operators, called annihilation operator and creation operator respectively.

$$\begin{aligned} a &= \frac{1}{\sqrt{2}}(Q + iP) \ , \\ a^\dagger &= \frac{1}{\sqrt{2}}(Q - iP) \ . \end{aligned} \tag{9.2.8}$$

In terms of these the operators x and p can be written as

$$\begin{aligned} x &= \sqrt{\frac{\hbar}{2m\omega}}(a^\dagger + a) \ , \\ p &= i\sqrt{\frac{m\hbar\omega}{2}}(a^\dagger - a) \ . \end{aligned} \tag{9.2.9}$$

Recall that P and Q are both self-adjoint on our Hilbert space which is $\mathcal{L}_2(-\infty, \infty)$. Now consider the commutator

$$\begin{aligned} [a, a^\dagger] &= \frac{1}{2}[Q + iP, Q - iP] \\ &= \frac{1}{2}\{[Q, Q] + i[P, Q] - i[Q, P] + [P, P]\} \\ &= \frac{1}{2}[0 + 1 + 1 + 0] = 1 \ . \end{aligned} \tag{9.2.10}$$

Thus,

$$[a, a^\dagger] = 1 \ . \tag{9.2.11}$$

Solving (9.2.8) we get,

$$\begin{aligned} Q &= \frac{1}{\sqrt{2}}(a^\dagger + a) \ , \\ P &= \frac{i}{\sqrt{2}}(a^\dagger - a) \ . \end{aligned} \tag{9.2.12}$$

Therefore,

$$\begin{aligned} H &= \frac{1}{2}\hbar\omega\left(Q^2+P^2\right) \\ &= \frac{1}{4}\hbar\omega[(a^\dagger+a)^2-(a^\dagger-a)^2] \\ &= \frac{1}{4}\hbar\omega[(a^\dagger)^2+a^\dagger a+aa^\dagger+a^2-(a^\dagger)^2+a^\dagger a+aa^\dagger-a^2] \end{aligned} \tag{9.2.13}$$

or

$$H=\frac{1}{2}\hbar\omega(a^\dagger a+aa^\dagger) \ . \tag{9.2.14}$$

Using (9.2.11) we can rewrite the Hamiltonian H to read

$$H=\hbar\omega(a^\dagger a+\frac{1}{2})=\hbar\omega(aa^\dagger-\frac{1}{2}) \ . \tag{9.2.15}$$

Furthermore,

$$[H,a]=-\hbar\omega\, a \tag{9.2.16}$$

and

$$[H,a^\dagger]=\hbar\omega\, a^\dagger \ . \tag{9.2.17}$$

Now since P and Q are self-adjoint it follows from (9.2.6) that H is a nonnegative self-adjoint operator. Thus, all the eigenvalues of H are nonnegative. In fact they are positive.

So, let ψ_0 be the lowest eigenstate of H. That is, ψ_0 is the state corresponding to the smallest eigenvalue. Let this eigenvalue be E_0. Then,

$$H\ \psi_0=E_0\ \psi_0 \ . \tag{9.2.18}$$

And hence

$$aH\ \psi_0=E_0\ a\psi_0 \ . \tag{9.2.19}$$

But,

$$\begin{aligned} aH &= Ha+[a,H] \\ &= Ha+\hbar\omega \ . \end{aligned} \tag{9.2.20}$$

Therefore, we find

$$(Ha+\hbar\omega a)\ \psi_0=E_0\ (a\psi_0) \tag{9.2.21}$$

and hence

$$H(a\ \psi_0)=(E_0-\hbar\omega)\ a\psi_0 \ . \tag{9.2.22}$$

Thus, clearly we have that either

$$a\psi_0=0 \tag{9.2.23}$$

or $a\psi_0$ is another eigenfunction of H corresponding to the eigenvalue $E_0 - \hbar\omega$. On the other hand, E_0 is the lowest eigenvalue by assumption. Thus,

$$a\psi_0 = 0 \ .$$

From (9.2.23) we also get, by multiplying by $\hbar\omega a^\dagger$ from the left, the equation

$$\hbar\omega a^\dagger a \ \psi_0 = 0 \ . \tag{9.2.24}$$

But,

$$\hbar\omega a^\dagger a \ = H - \frac{1}{2}\hbar\omega \ . \tag{9.2.25}$$

Therefore,

$$(H - \frac{1}{2}\hbar\omega)\psi_0 = 0 \tag{9.2.26}$$

or

$$H \ \psi_0 = \frac{1}{2}\hbar\omega \ \psi_0 \ . \tag{9.2.27}$$

So we have found the lowest eigenvalue, namely

$$E_0 = \frac{1}{2}\hbar\omega \ . \tag{9.2.28}$$

Now we operate on (9.2.27) once more with $a^\dagger$ to get:

$$a^\dagger H \ \psi_0 = \frac{1}{2}\hbar\omega \ a^\dagger\psi_0 \ . \tag{9.2.29}$$

Using (9.2.17) again this gives

$$(Ha^\dagger - \hbar\omega a^\dagger) \ \psi_0 = \frac{1}{2}\hbar\omega \ a^\dagger\psi_0 \ . \tag{9.2.30}$$

Therefore,

$$H \ (a^\dagger\psi_0) = (1 + \frac{1}{2})\hbar\omega \ (a^\dagger\psi_0) \ . \tag{9.2.31}$$

So $a^\dagger\psi_0$ is another eigenfunction corresponding to the eigenvalue $(1+\frac{1}{2})\hbar\omega$. This shows that $a^\dagger$ is a “raising operator”; it raises the eigenvalue of the Hamiltonian by one unit. Hence it is called a raising or step-up operator . We also call it a *creation operator* since it creates a quantum of energy $\hbar\omega$. If we repeat this procedure n times we get

$$H \ \left((a^\dagger)^n\psi_0\right) = (n + \frac{1}{2})\hbar\omega \ \left((a^\dagger)^n\psi_0\right) \ . \tag{9.2.32}$$

Thus, in this manner we get a whole sequence of eigenfunctions. We now find $\psi_0(x)$ explicitly. We have

$$a\psi_0(x) = 0 \ . \tag{9.2.33}$$

Writing this out we get

$$(Q+iP)\psi_0(x)=0 \tag{9.2.34}$$

or

$$\left(\sqrt{\frac{m\omega}{\hbar}}\,x+\sqrt{\frac{\hbar}{m\omega}}\,\frac{d}{dx}\right)\psi_0(x)=0 \tag{9.2.35}$$

which simplifies to

$$\left(\frac{d}{dx}+\frac{k}{\hbar\omega}\right)\psi_0(x)=0\,. \tag{9.2.36}$$

The solution of this equation is

$$\psi_0(x)=A\,e^{-kx^2/2\hbar\omega}\,. \tag{9.2.37}$$

To normalize ψ_0 we compute

$$\begin{aligned}(\psi_0,\psi_0)&=1\\&=|A|^2\int_{-\infty}^{\infty}e^{-kx^2/\hbar\omega}\,dx\\&=|A|^2\sqrt{\frac{\pi\hbar\omega}{k}}\,.\end{aligned} \tag{9.2.38}$$

Thus, choosing the phase of A we get

$$A=\left(\frac{k}{\pi\hbar\omega}\right)^{1/4} \tag{9.2.39}$$

and hence

$$\psi_0(x)=\left(\frac{k}{\pi\hbar\omega}\right)^{1/4}e^{-kx^2/2\hbar\omega}\,. \tag{9.2.40}$$

Furthermore, note that this solution is unique. Why do we emphasize this? The reason is that it allows us later to conclude that we have found all the eigenfunctions. First, however, we shall find the proper normalization for all the $\psi_n(x)$. Thus, we have

$$\psi_n(x)=c_n\left(a^\dagger\right)^n\psi_0(x)\,. \tag{9.2.41}$$

The factor c_n is included for normalization, that is, to make

$$(\psi_n,\psi_n)=1\,. \tag{9.2.42}$$

Then,

$$\begin{aligned}
(\psi_n, \psi_n) &= 1 \\
&= \left|\frac{c_n}{c_{n-1}}\right|^2 (a^\dagger \psi_{n-1}, a^\dagger \psi_{n-1}) \\
&= \left|\frac{c_n}{c_{n-1}}\right|^2 (\psi_{n-1}, a a^\dagger \psi_{n-1}) \\
&= \left|\frac{c_n}{c_{n-1}}\right|^2 \frac{1}{\hbar\omega}(\psi_{n-1}, (H + 1/2\hbar\omega)\psi_{n-1}) \\
&= \left|\frac{c_n}{c_{n-1}}\right|^2 n(\psi_{n-1}, \psi_{n-1}) \; . \qquad (9.2.43)
\end{aligned}$$

Therefore,

$$|c_n|^2 = \frac{1}{n}|c_{n-1}|^2 \; . \qquad (9.2.44)$$

Using $c_0 = 1$ we can solve this recursion relation to get

$$c_n = \frac{1}{\sqrt{n!}} \qquad (9.2.45)$$

where we have made an arbitrary choice of phase. Thus,

$$\psi_n(x) = \frac{1}{\sqrt{n!}} \left[\sqrt{\frac{m\omega}{2\hbar}}\, x - \sqrt{\frac{\hbar}{2m\omega}} \frac{d}{dx}\right]^n \psi_0(x) \qquad (9.2.46)$$

or

$$\psi_n(x) = (-1)^n \frac{1}{\sqrt{n!}} \left(\frac{\hbar\omega}{2k}\right)^{n/2} \left[\frac{d}{dx} - \frac{k}{\hbar\omega} x\right]^n \left(\frac{k}{\pi\hbar\omega}\right)^{1/4} e^{-kx^2/2\hbar\omega} \; . \qquad (9.2.47)$$

These eigenfunctions are the *hermite functions.* We already know that they form a complete orthonormal set. Thus, since $\psi_0(x)$ is unique, we have all the eigenfunctions of H. It is a simple matter to check the orthogonality. Consider first

$$\begin{aligned}
(\psi_0, \psi_1) &= c_1(\psi_0, a^\dagger \psi_0) \\
&= c_1(a\psi_0, \psi_0) = 0 \qquad (9.2.48)
\end{aligned}$$

since $a\psi_0 = 0$. Similarly

$$\begin{aligned}
(\psi_0, \psi_n) &= c_n(\psi_0, (a^\dagger)^n \psi_0) \\
&= c_1(a\psi_0, (a^\dagger)^{n-1}\psi_0) = 0 \; . \qquad (9.2.49)
\end{aligned}$$

If we now consider (ψ_n, ψ_m) with $n < m$ we are led to consider expressions of the form

$$\begin{aligned}
((a^\dagger)^n \psi_0, (a^\dagger)^m \psi_n) &= n!(a a^\dagger (a^\dagger)^{n-1}\psi_0, (a^\dagger)^{m-1}\psi_0) \\
&= \frac{1}{\hbar\omega}((H + 1/2\,\hbar\omega)(a^\dagger)^{n-1}\psi_0, (a^\dagger)^{m-1}\psi_0) \\
&= n((a^\dagger)^{n-1}\psi_0, (a^\dagger)^{m-1}\psi_0) \; . \qquad (9.2.50)
\end{aligned}$$

Repeating this we get eventually

$$\begin{aligned}((a^\dagger)^n\psi_0, (a^\dagger)^m\psi_n) &= n!(\psi_0, (a^\dagger)^{m-n}\psi_0) \\ &= n!(a\psi_0, (a^\dagger)^{n-m-1}\psi_0) = 0 \ . \end{aligned} \tag{9.2.51}$$

Notice that throughout these computations the crucial formula was

$$a\psi_0 = 0 \ . \tag{9.2.52}$$

Thus we have solved the simple harmonic oscillator using algebraic techniques except to find ψ_0 explicitly. Notice, however, that we did not need to find ψ_0 explicitly. In fact all the results that we obtained followed directly from the operator properties of H, p and x. The usefulness of a and $a^\dagger$ do not end here. In fact all physical quantities pertaining to the simple harmonic oscillator may be calculated using only a and $a^\dagger$. Later we shall have occasion to re-examine the time-dependence of the simple harmonic oscillator. We shall then find that a and $a^\dagger$ also have analogous quantities classically, and these classical quantities simplify the classical computation. Now, however, we illustrate some of the applications of a and $a^\dagger$.

9.2.1 Expectation Values

Earlier we expressed the position and momentum operatores, x and p in terms of the creation and annihilation operators $a^\dagger$ and a

$$\begin{aligned} x &= \sqrt{\frac{\hbar}{2m\omega}}(a^\dagger + a) \\ p &= i\sqrt{\frac{m\hbar\omega}{2}}(a^\dagger - a) \ . \end{aligned} \tag{9.2.53}$$

Thus, we can evaluate expectation values of x and p by using only a and $a^\dagger$. For example

$$\begin{aligned} (\psi_n, p\psi_n) &= \sqrt{m\hbar\omega}(\psi_n, P\psi_n) \\ &= i\sqrt{m\hbar\omega/2}\left[(\psi_n, a^\dagger\psi_n) - (\psi_n, a\psi_n)\right] \\ &= 0 \ . \end{aligned} \tag{9.2.54}$$

Similarly,

$$(\psi_n, x\psi_n) = 0 \ .$$

This is to be expected since the average momentum and position of a simple harmonic oscillator are indeed zero. The same applies for any other dynamical variable. For example, consider the kinetic energy

$$\begin{aligned} T &= \frac{p^2}{2m} \\ &= \frac{m\hbar\omega}{2m}P^2 \\ &= \frac{1}{2}\hbar\omega(-\frac{1}{2})(a^\dagger - a)(a^\dagger - a) \end{aligned} \tag{9.2.55}$$

or multiplying out

$$\begin{aligned} T &= \frac{1}{4}(aa^\dagger + a^\dagger a - a^{\dagger 2} - a^2) \\ &= \frac{1}{2}H - \frac{1}{4}\hbar\omega(a^{\dagger 2} + a^2)\,. \end{aligned} \tag{9.2.56}$$

Thus,

$$\begin{aligned} (\psi_n, T\psi_n) &= \frac{1}{2}(\psi_n, H\psi_n) - \frac{1}{4}\hbar\omega(\psi_n, (a^{\dagger 2} + a^2)\psi_n) \\ &= \frac{1}{2}(\psi_n, H\psi_n) \\ &= \frac{1}{2}(n + \frac{1}{2})\,\hbar\omega\ . \end{aligned} \tag{9.2.57}$$

Actually we have proven more since we showed that

$$\langle T\rangle = \frac{1}{2}\langle H\rangle\,. \tag{9.2.58}$$

This is the same as the classical virial theorem and states that the average kinetic energy equals one-half the average total energy. Also we have

$$\langle V\rangle = \langle H\rangle - \langle T\rangle = \frac{1}{2}\langle H\rangle\,. \tag{9.2.59}$$

Another set of useful formulae is obtained using (9.2.41) and (9.2.44). Thus we have:

$$a^\dagger\psi_n = \frac{c_n}{c_{n+1}}\psi_{n+1} = \sqrt{n+1}\psi_{n+1}\ , \tag{9.2.60}$$

showing clearly why $a^\dagger$ is called a "step-up" operator. It also shows why $a^\dagger$ is known as a "creation" operator since it "creates" a higher eigenstate. Applying a to this equation we get

$$aa^\dagger\psi_n = \sqrt{n+1}a\psi_{n+1}\ . \tag{9.2.61}$$

Therefore,

$$\begin{aligned} a\psi_n &= \frac{1}{\sqrt{n}}aa^\dagger\psi_{n-1} \\ &= \frac{1}{\sqrt{n}}\frac{1}{\hbar\omega}(H + 1/2\hbar\omega))\psi_{n-1}\ . \end{aligned} \tag{9.2.62}$$

Thus,

$$a\psi_n = \sqrt{n}\psi_{n-1}\ . \tag{9.2.63}$$

So a is a "lowering" or "step-down" operator. It is also sometimes known as an "annihilation" operator. To further show how these operators are used we consider a matrix element of the form

$$\begin{aligned} (\psi_r, p\psi_s) &= i\sqrt{m\hbar\omega/2}\left[(\psi_r, a^\dagger\psi_s) - (\psi_r, a\psi_s)\right] \\ &= i\sqrt{m\hbar\omega/2}\left[\sqrt{s+1}(\psi_r, \psi_{s+1}) - \sqrt{s}(\psi_r, \psi_{s-1})\right]\ . \end{aligned} \tag{9.2.64}$$

Thus, we have found that

$$(\psi_r, p\psi_s) = i\sqrt{\frac{\hbar}{2m\omega}}\left[\sqrt{s+1}\delta_{r,s+1} - \sqrt{s}\delta_{r,s-1}\right] \quad . \tag{9.2.65}$$

In a similar manner we get

$$(\psi_r, x\psi_s) = i\sqrt{m\hbar\omega/2}\left[\sqrt{s+1}\delta_{r,s+1} - \sqrt{s}\delta_{r,s-1}\right] \quad . \tag{9.2.66}$$

In this way it is possible to evaluate, in a purely algebraic manner, all matrix elements between simple harmonic oscillator eigenstates. We now turn to another problem that can also be handled by purely algebraic methods.

9.3 The Rigid Rotator

Consider a diatomic system which we approximate as follows. It is a "dumbbell" with masses attached at the ends of a rigid rod and free to rotate in any direction (figure 9.1). The moment of inertia about the line through the masses is assumed to be negligible (zero). If we assume that the center of mass of the system is fixed then the only motion possible is a rotation. In that case the classical Hamiltonian is

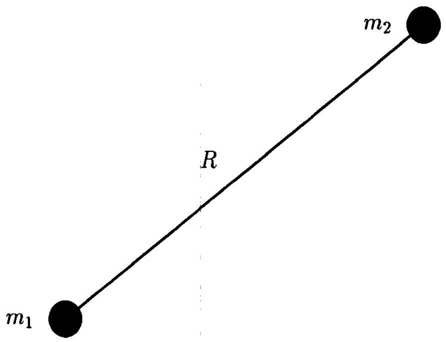

Figure 9.1: A dumbbell molecule.

$$H = \frac{\mathbf{L}^2}{2I} \tag{9.3.67}$$

where $\mathbf{L}$ is the total angular momentum and I is the moment of inertia about the axis through the center of mass and normal to the line connecting the masses. We wish to solve this problem quantum mechanically. Now,

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \, . \tag{9.3.68}$$

Thus,

$$\begin{aligned} L_x &= yp_z - zp_y \\ L_y &= zp_x - xp_z \\ L_z &= xp_y - yp_x \,. \end{aligned} \tag{9.3.69}$$

To quantize this system we simply replace $\mathbf{p}$ by $\hbar/i\ \nabla$. So the quantum mechanical angular momentum operator is given by

$$\mathbf{L} = \frac{\hbar}{i}\mathbf{r} \times \nabla \tag{9.3.70}$$

or

$$\begin{aligned} L_x &= i\hbar\left(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\right) \\ L_y &= i\hbar\left(x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}\right) \\ L_z &= i\hbar\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right) \,. \end{aligned} \tag{9.3.71}$$

Now introducing spherical coordinates

$$\begin{aligned} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \end{aligned} \tag{9.3.72}$$

we get

$$\begin{aligned} L_x &= i\hbar\left(\sin\theta\,\frac{\partial}{\partial\theta} + \cos\varphi\cot\theta\frac{\partial}{\partial\varphi}\right) \\ L_y &= i\hbar\left(\sin\varphi\cot\theta\,\frac{\partial}{\partial\varphi} - \cos\varphi\frac{\partial}{\partial\theta}\right) \\ L_z &= -i\hbar\frac{\partial}{\partial\varphi} \,. \end{aligned} \tag{9.3.73}$$

At this stage we can introduce a simplifying assumption in our problem. We assume that our "molecule" is constrained to rotations in a plane which we choose to be the $x-y$ plane. In that case we are considering a *One-Dimensional Rotator*. Furthermore, our Hamiltonian simplifies to

$$H = \frac{L_z^2}{2I} \,. \tag{9.3.74}$$

Writing out the eigenvalue problem we get

$$-\frac{\hbar^2}{2I}\frac{d^2\psi}{d\varphi^2} = E\psi \,. \tag{9.3.75}$$

Before proceeding we must consider whether L_z is self-adjoint. Thus, we look at the equation for the deficiency indices

$$-i\hbar \frac{df}{d\varphi} = \pm i\hbar f \; . \tag{9.3.76}$$

The solutions are

$$f_\pm = e^{\mp\varphi} \; . \tag{9.3.77}$$

Since $0 \leq \varphi \leq 2\pi$ both solutions are admissible and the deficiency indices are $(1,1)$. Thus we have a one-parameter family of self-adjoint extensions and we must pick an appropriate one. We do this partly by symmetry considerations. Thus, suppose $f(\varphi)$ is some physical state. In that case we want $f(\varphi + 2\pi)$ to represent the same state. This means that we must have

$$f(\varphi + 2\pi) = e^{i\alpha} f(\varphi) \; . \tag{9.3.78}$$

Here α is the parameter labelling the different self-adjoint extensions. What does this mean? It is easy to show (in the same manner we used to show that p_z is the generator of translations in the z-direction) that L_z is the generator of the translations of φ ; that is, L_z is the generator of rotations about the z-axis. Thus, rotating by 2π about the z-axis brings us back to the original point and we must specify the boundary condition. The usual argument given is that the wave function should be single-valued and hence $\alpha = 0$. This argument, although correct, is not completely satisfactory since none of our assumptions about quantum mechanics, made so far, required the wave function to be single-valued. We now summarize the results of a careful investigation [9.1] that shows why the wave-function must indeed be single-valued.

For a spherically symmetric system, it is physically necessary that not only should one component of the angular momentum (L_z in this case) be conserved but that, in fact, the total angular momentum

$$L^2 = L_x^2 + L_y^2 + L_z^2 \tag{9.3.79}$$

must be conserved. Thus, we must be able to simultaneously diagonalize L^2 and some one component say

$$L_n = \hat{n}.\mathbf{L} \tag{9.3.80}$$

of the angular momentum. Here, $\hat{n}$ is a unit vector. This requires

$$[L_n, L^2] = 0 \; . \tag{9.3.81}$$

The above relation follows formally if one works out the commutation relations between L_x , L_y , L_z and L^2. We shall do this later. Furthermore, since we can arbitrarily decide how we shall label a set of three orthogonal axes as x , y and z, the three operators L_x , L_y , L_z must be equivalent. Consider now the self-adjoint extensions L_z^α corresponding to the boundary conditions (9.3.78). In this case there are corresponding self-adjoint extensions for L_x and L_y say L_x^α and L_y^α . The results for these extensions are as follows.

1) The eigenvalues of L_x^α and L_y^α are integral regardless of α. The eigenvalues of L_z^α are integral only for $\alpha = 0$.

2) For $\alpha = 0$

$$L^2\psi = L_x^2\psi + L_y^2\psi + L_z^2\psi$$

for every $\psi \in D_{L^2}$, where we have also dropped the superscript. For $\alpha \neq 0$ this relation is false.

3) If $\alpha = 0$

$$[L^2, L_x^\alpha] = [L^2, L_y^\alpha] = [L^2, L_z^\alpha] = 0 \ .$$

If $\alpha \neq 0$ we still have

$$[L^2, L_z^\alpha] = 0$$

but L^2 does not commute with L_x^α and L_y^α. Thus, neither L^2 and L_x^α nor L^2 and L_y^α have simultaneous eigenfunctions for $\alpha \neq 0$.

4) If $\alpha = 0$ and $\psi \in D_{L^2}$ then

$$L_x L_y \psi - L_y L_x \psi = i\hbar L_z \psi \ .$$

For $\alpha \neq 0$ this is no longer true. In this latter case there are eigenfunctions of L^2 to which neither $L_x L_y$ nor $L_y L_x$ can be applied.

These results show that for $\alpha \neq 0$ not only is there a special preferred direction, the z-direction, but even the algebra of angular momentum operators no longer holds. However, this is not the case physically. Thus, the only *physically acceptable* self-adjoint extension is given by $\alpha = 0$. This corresponds to having single-valued wavefunctions. For this case we drop the superscript. Notice, that mathematics alone could not decide on the appropriate self-adjoint extension, the decision that the wavefunction has to be single-valued had to be made on physical grounds. For more details on this question we list a series of references at the end of this chapter. We can now return to a consideration of our original problem of the one-dimensional rotator

$$H = \frac{{L_z}^2}{2I} = -\frac{\hbar^2}{2I}\frac{d^2}{d\varphi^2} \tag{9.3.82}$$

where $f \in D_{L_z}$ requires that

$$f(\varphi + 2\pi) = f(\varphi) \tag{9.3.83}$$

and for $f \in D_{L_z^2}$ also that f has to be twice differentiable. Clearly,

$$[L_z, H] = 0 \ . \tag{9.3.84}$$

Thus, we can simultaneously diagonalize L_z and H. The eigenvalue problem for L_z can be written

$$-i\hbar\frac{d\Phi}{d\varphi} = m\hbar\Phi \ . \tag{9.3.85}$$

Thus,

$$\Phi_m(\varphi) = A\,e^{im\varphi} \ . \tag{9.3.86}$$

Applying the boundary condition with $\alpha = 0$ we get

$$e^{im(\varphi+2\pi)} = e^{im\varphi} \ . \tag{9.3.87}$$

Thus,

$$e^{im2\pi} = 1 \ . \tag{9.3.88}$$

Hence

$$m = 0, \pm 1, \pm 2, \dots \ . \tag{9.3.89}$$

If we use the Hamiltonian H we get

$$H\Phi_m = E_m\Phi_m = \frac{L_z^2}{2I}\Phi_m \tag{9.3.90}$$

and this reads

$$E_m\Phi_m = \frac{\hbar^2}{2I}m^2\Phi_m \tag{9.3.91}$$

which shows that the energy eigenvalues are

$$E_m = \frac{\hbar^2}{2I}m^2 \quad m = 0, \pm 1, \pm 2, \dots \ . \tag{9.3.92}$$

The normalization is obtained from

$$\int_0^{2\pi} |\Phi_m|^2 \, d\varphi = |A|^2 \int_0^{2\pi} d\varphi = 1 \ . \tag{9.3.93}$$

Thus, choosing the phase of A we get

$$A = \frac{1}{\sqrt{2\pi}} \ . \tag{9.3.94}$$

Hence,

$$\Phi_m = \frac{1}{\sqrt{2\pi}}e^{im\varphi} \ . \tag{9.3.95}$$

9.4 3D Rigid Rotator: Angular Momentum

We are now ready for the full rigid rotator problem. Here

$$H = \frac{\mathbf{L}^2}{2I} = \frac{1}{2I}\left(L_x^2 + L_y^2 + L_z^2\right) \ . \tag{9.4.96}$$

Clearly,

$$[\mathbf{L}^2, H] = 0 \tag{9.4.97}$$

and as we show next, the projection of **L** along any direction $\hat{n}$ commutes with L^2.

$$[L_n, \mathbf{L}^2] = [L_n, H] = 0 \ . \tag{9.4.98}$$

We use the form (9.3.68) for **L** to compute several commutators. Now,

$$\begin{aligned} [L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] \\ &= [yp_z, zp_x] - [zp_y, zp_x] - [yp_z, xp_z] + zp_y, xp_z] \ . \end{aligned} \tag{9.4.99}$$

Consider the first term

$$\begin{aligned} [yp_z, zp_x] &= y[p_z, zp_x] + [y, zp_x]p_z \\ &= y[p_z, z]p_x + yz[p_z, p_x] + [y, z]p_xp_z + z[y, p_x]p_z \\ &= -i\hbar yp_x + 0 + 0 + 0 = -i\hbar yp_x \ . \end{aligned} \tag{9.4.100}$$

The second and third terms both contribute nothing. The last term contributes $i\hbar xp_y$. Thus,

$$[L_x, L_y] = i\hbar L_z \ . \tag{9.4.101}$$

Proceeding in this manner we get the commutation relations

$$[L_y, L_z] = i\hbar L_x \ , \tag{9.4.102}$$

$$[L_z, L_x] = i\hbar L_y \ . \tag{9.4.103}$$

Notice that these are just cyclic permutations of x, y, z as shown.

$$\begin{tikzcd} & y \arrow[dr] & \\ x \arrow[ur] & & z \arrow[ll] \end{tikzcd}$$

We further compute

$$\begin{aligned} [L_x, L^2] &= [L_x, L^2] + [L_x, L_y^2] + [L_x, L_z^2] \\ &= 0 + L_y[L_x, L_y] + [L_x, L_y]L_y + L_z[L_x, L_z] + [L_x, L_z]L_z \\ &= i\hbar(L_yL_z + L_zL_y - L_yL_z - L_zL_y) = 0 \ . \end{aligned} \tag{9.4.104}$$

Thus,

$$[L_x, L^2] = 0 \ . \tag{9.4.105}$$

Similarly,

$$[L_y, L^2] = 0 \tag{9.4.106}$$

and

$$[L_z, L^2] = 0 \, . \tag{9.4.107}$$

We have, therefore, shown that L^2 and one component of **L** can be simultaneously diagonalized. Since L_z has the simplest representation as a differential operator (in spherical coordinates) one chooses L_z.

Thus, calling the eigenvalues of L^2, $l(l+1)\hbar^2$ and of L_z, $m\hbar$ we have to solve the following eigenvalue problems

$$L^2 Y_{lm}(\theta, \varphi) = l(l+1)\hbar^2 Y_{lm}(\theta, \varphi) \tag{9.4.108}$$

$$L_z Y_{lm}(\theta, \varphi) = m\hbar Y_{lm}(\theta, \varphi) \, . \tag{9.4.109}$$

The energy eigenvalues are then given by

$$E_l = \frac{\hbar^2}{2I} l(l+1) \, . \tag{9.4.110}$$

The factors $l(l+1)\hbar^2$ are chosen for later convenience. We could obviously have chosen anything else and then wound up with this result. If we now use (9.3.73) and the definition of L^2 we find

$$L^2 = -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right] \, . \tag{9.4.111}$$

This form becomes much more significant if we look at the Laplace operator in spherical coordinates

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{L^2}{\hbar^2 r^2} \, . \tag{9.4.112}$$

Thus, the angular portion of ∇^2 is given completely by L^2.

In order to solve (9.4.108), (9.4.109) we try to separate variables. Therefore, we set

$$Y_{lm}(\theta, \varphi) = c_{lm} P_l^m(\cos\theta) \Phi_m(\varphi) \tag{9.4.113}$$

where c_{lm} is a normalization constant. Then,

$$Y_{lm}(\theta, \varphi) = \frac{c_{lm}}{\sqrt{2\pi}} e^{im\varphi} P_l^m(\cos\theta) \, . \tag{9.4.114}$$

So that (9.4.109) is automatically satisfied and (9.4.108) becomes

$$\left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) - \frac{m^2}{\sin^2\theta} \right] P_l^m(\cos\theta) = -l(l+1) P_l^m(\cos\theta) \, . \tag{9.4.115}$$

If we set $\cos\theta = x$ then

$$\frac{d}{dx} = -\frac{1}{\sin\theta}\frac{d}{d\theta} . \tag{9.4.116}$$

Writing out the expression (9.4.115) it reads

$$\frac{d}{dx}\left[(1-x^2)\frac{dP_l^m}{dx}\right] + \left[l(l+1) - \frac{m^2}{1-x^2}\right]P_l^m = 0 . \tag{9.4.117}$$

Simplifying this equation we obtain

$$(1-x^2)\frac{d^2P}{dx^2} - 2x\frac{dP}{dx} + \left[l(l+1) - \frac{m^2}{1-x^2}\right]P = 0 . \tag{9.4.118}$$

This is the differential equation for the *associated Legendre functions* which are denoted by $P_l^m(x)$. For $m = 0$ it is the differential equation for the Legendre polynomials $P_l(x)$. Rather than solve the differential equation (9.4.118) we shall proceed as for the simple harmonic oscillator and solve the problem algebraically. Since in the algebraic solution we are not restricted to having a representation of the angular momentum operator **L** as a differential operator in ordinary space we shall find not only the solutions above but also some new solutions that do not correspond to orbital angular momentum but to something new called spin angular momentum.

9.5 Algebraic Approach to Angular Momentum

The assumptions involved are as follows.
1) Commutation Relations

$$[L_x, L_y] = i\hbar L_z$$

and cyclic permutations.
2) Angular momentum is an observable. This means that every component of **L** is a self adjoint operator. It then follows that

$$L^2 = L_x^2 + L_y^2 + L_z^2$$

is a positive, self-adjoint operator. Thus,

$$L^2 \geq 0 \quad \text{and} \quad L^2 \geq L_z^2 \ . \tag{9.5.119}$$

Also as we saw before

$$[\mathbf{L}, L^2] = 0 . \tag{9.5.120}$$

Thus, we may diagonalize L^2 and one component of **L** simultaneously. It is conventional to choose L_z. We now introduce two non-hermitian operators $L_\pm$

$$L_\pm = L_x \pm iL_y \tag{9.5.121}$$

analogous to the raising or creation and lowering or annihilation operators $a^\dagger$ and a used with the simple harmonic oscillator. We also need the commutation relations for $L_\pm$. First we note that it follows directly from (9.5.120) that

$$[L_\pm, L^2] = 0 \ . \tag{9.5.122}$$

Also,

$$\begin{aligned} [L_z, L_\pm] &= [L_z, L_x] \pm i[L_z, L_y] \\ &= i\hbar L_y \pm i(-i\hbar L_x) \\ &= \pm\hbar(L_x \pm iL_y) \ . \end{aligned} \tag{9.5.123}$$

Hence,

$$[L_z, L_\pm] = \pm\hbar L_\pm \ . \tag{9.5.124}$$

Finally we also get by straightforward multiplication that

$$L_\pm L_\mp = L_x^2 + L_y^2 \pm \hbar L_z \tag{9.5.125}$$

so that

$$L_\pm L_\mp = L^2 - L_z^2 \pm \hbar L_z \ . \tag{9.5.126}$$

We now consider the simultaneous eigenvalue problems

$$L^2\psi = a\hbar^2\psi \tag{9.5.127}$$

$$L_z\psi = b\hbar\psi \ . \tag{9.5.128}$$

Since

$$L^2 \geq L_z^2 \tag{9.5.129}$$

we get that

$$a\hbar^2 \geq b^2\hbar^2 \tag{9.5.130}$$

or

$$a \geq b^2 \ . \tag{9.5.131}$$

Now apply L_+ to (9.5.128) and use (9.5.124). This gives

$$L_+L_z\psi = b\hbar L_z\psi = L_zL_+\psi - \hbar L_+\psi \ . \tag{9.5.132}$$

Thus,

$$L_z(L_+\psi) = (b+1)\hbar(L_+\psi) \ . \tag{9.5.133}$$

This shows that $L_+\psi$ is a new eigenfunction of L_z with eigenvalue $(b+1)\hbar$. Since $L_\pm$ and L^2 commute, $L_+\psi$ is still an eigenfunction of L^2 with the eigenvalue $a\hbar^2$. However due to (9.5.131) we see that there must exist a maximum eigenvalue for L_z for otherwise we could violate (9.5.131). Suppose ψ_{max} is this eigenfunction. In that case applying L_+ yields (9.5.133). Thus, consistency can only be maintained if

$$L_+\psi_{max} = 0 \tag{9.5.134}$$

where ψ_{max} is the eigenfunction corresponding to the *largest eigenvalue* of L_z. Now we apply L_- to (9.5.134) and use (9.5.126). Then we get

$$L_-L_+\psi_{max} = (L^2 - L_z^2 - \hbar L_z)\psi_{max} = 0 \tag{9.5.135}$$

or

$$(a\hbar^2 - b^2\hbar^2 - b\hbar^2)\psi_{max} = 0 \,. \tag{9.5.136}$$

Thus,

$$a = b^2 + b \,. \tag{9.5.137}$$

We now proceed downwards from the largest eigenvalue. Thus, we operate on (9.5.128) with L_- and use (9.5.124) to get

$$L_z(L_-\psi_{max}) = (b-1)\hbar(L_-\psi_{max}) \,. \tag{9.5.138}$$

Repeating this n times we get

$$L_z(L_-^n\psi_{max}) = (b-n)\hbar(L_-^n\psi_{max}) \,. \tag{9.5.139}$$

Now by making n large enough we can again violate (9.5.131). This means we must reach a point where

$$\psi' = (L_-^n\psi_{max}) \tag{9.5.140}$$

and

$$L_-\psi' = 0 \,. \tag{9.5.141}$$

Applying L_+ to this equation and using (9.5.126) then gives:

$$L_+L_-\psi' = (L^2 - L_z^2 - \hbar L_z)(L_-^n\psi_{max}) = 0 \tag{9.5.142}$$

or

$$[a\hbar^2 - (b-n)^2\hbar^2 + (b-n)\hbar^2](L_-^n\psi_{max}) = 0 \ . \tag{9.5.143}$$

Thus,

$$a = (b-n)^2 - (b-n) \ . \tag{9.5.144}$$

We combine this with (9.5.137) and get:

$$2b(n+1) = n(n+1) \,. \tag{9.5.145}$$

Thus,

$$b = \frac{n}{2} \equiv l \tag{9.5.146}$$

where

$$l = 0\,,\ 1/2\,,\ 1\,,\ 3/2\,,\ 2\ 5/2\,,\ \ldots\ .$$

Also,

$$a = b(b+1) = l(l+1)\ . \tag{9.5.147}$$

So, we can label the eigenstates with l and m such that

$$L^2\psi_{lm} = l(l+1)\hbar^2\psi_{lm} \tag{9.5.148}$$

$$L_z\psi_{lm} = m\hbar\psi_{lm} \tag{9.5.149}$$

where l is the largest value of $|m|$.

So $l \geq |m|$ or

$$m = l, l-1, l-2, \ldots, -(l-1), -l\ . \tag{9.5.150}$$

As we have already seen, for the orbital angular momentum l is an integer, and the ψ_{lm} are the spherical harmonics $Y_{l,m}(\theta,\varphi)$. We now re-examine them using the algebraic techniques. The case of half-odd-integer angular momentum is examined in section 9.7. Consider

$$\begin{aligned} L_z &= \frac{\hbar}{i}\frac{\partial}{\partial\varphi} \\ L_\pm &= \hbar\, e^{i\varphi}\left[\pm\frac{\partial}{\partial\theta} + i\cot\theta\,\frac{\partial}{\partial\varphi}\right] \\ L^2 &= -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]\ . \end{aligned} \tag{9.5.151}$$

If we start with the state of lowest angular momentum, $Y_{l,-l}(\theta,\varphi)$ we have

$$L^2 Y_{l,-l} = l(l+1)\hbar^2 Y_{l,-l} \tag{9.5.152}$$

$$L_z Y_{l,-l} = -l\hbar Y_{l,-l} \tag{9.5.153}$$

and

$$L_- Y_{l,-l} = 0\ . \tag{9.5.154}$$

We can also work up from this point using L_+. Thus,

$$Y_{l,-l+1} = cL_+ Y_{l,-l} \tag{9.5.155}$$

and so on, eventually reaching the highest wavefunction $Y_{l,l}$. Then,

$$\begin{aligned} L^2 Y_{l,l} &= l(l+1)\hbar^2 Y_{l,l} \\ L_z Y_{l,l} &= l\hbar Y_{l,l} \\ L_+ Y_{l,l} &= 0\ . \end{aligned} \tag{9.5.156}$$

To look for the eigenfunctions explicitly we proceed as before by writing

$$Y_{l,m} = \frac{1}{\sqrt{2\pi}} e^{im\varphi} P_l^m(\cos\theta) \ . \tag{9.5.157}$$

We then get from

$$L_+ Y_{l,l} = 0 \tag{9.5.158}$$

that

$$e^{i\varphi}\left[e^{il\varphi}\frac{dP_l^l}{d\theta} - l \cot\theta e^{il\varphi} P_l^l\right] = 0 \tag{9.5.159}$$

or

$$\frac{dP_l^l}{d\theta} = l \cot\theta P_l^l \ . \tag{9.5.160}$$

Thus,

$$P_l^l = A_l \sin^l\theta \ . \tag{9.5.161}$$

We get A_l by normalization

$$|A_l|^2 I_l = |A_l|^2 \int_0^\pi \sin^{2l}\theta \sin\theta \, d\theta = 1 \ . \tag{9.5.162}$$

This leads to the recursion relation (see problem 9.19)

$$I_l = \frac{2l}{2l+1} I_{l-1} \ . \tag{9.5.163}$$

Iterating this we find

$$\begin{aligned} I_l &= \frac{2l}{2l+1} I_{l-1} \\ &= \frac{2l}{2l+1} \cdot \frac{2(l-1)}{2(l-1)+1} I_{l-2} \\ &= \frac{2^2 \, l(l-1)}{(2l+1)(2l-1)} I_{l-2} \\ &= \frac{2^l \, l!}{(2l+1)(2l-1)\ldots 3.1} I_0 \ . \end{aligned} \tag{9.5.164}$$

This simplifies to

$$I_l = \frac{2^{2l+1}(l!)^2}{(2l+1)!} \ . \tag{9.5.165}$$

Choosing the phase of A_l to be $(-1)^l$ we get

$$P_l^l = (-1)^l \sqrt{\frac{(2l+1)!}{2}} \frac{1}{2^l \, l!} \sin^l\theta \ . \tag{9.5.166}$$

Thus,

$$Y_{l,l}(\theta,\varphi) = \frac{(-1)^l}{\sqrt{2\pi}}\sqrt{\frac{(2l+1)!}{2}}\frac{\sin^l\theta}{2^l\,l!}\sin\theta\, e^{il\varphi} \, . \tag{9.5.167}$$

To get the other spherical harmonics we apply L_-. So we find

$$Y_{l,l-1} = c_{l,l-1}\, L_- Y_{l,l} \tag{9.5.168}$$

and finally

$$Y_{l,m} = c_{l,m}\, L_-^{l-m} Y_{l,l} \, . \tag{9.5.169}$$

Therefore, we can write

$$Y_{l,m} = \frac{c_{l,m}}{c_{l,m+1}} L_- Y_{l,m+1} \, . \tag{9.5.170}$$

To evaluate the normalization constants $c_{l,m}$ we simply use

$$(Y_{l,m}, Y_{l,m}) = \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta\, d\theta\, Y_{l,m}^* Y_{l,m} = 1 \tag{9.5.171}$$

to get

$$\begin{aligned} 1 &= \left|\frac{c_{l,m}}{c_{l,m+1}}\right|^2 (L_- Y_{l,m+1},\, L_- Y_{l,m+1}) \\ &= \left|\frac{c_{l,m}}{c_{l,m+1}}\right|^2 (Y_{l,m+1},\, L_+ L_- Y_{l,m+1}) \\ &= \left|\frac{c_{l,m}}{c_{l,m+1}}\right|^2 \left(Y_{l,m+1},\, (L^2 - L_z^2 + \hbar L_z) Y_{l,m+1}\right) \\ &= \left|\frac{c_{l,m}}{c_{l,m+1}}\right|^2 \hbar^2 [l(l+1) - m(m+1)] \, . \end{aligned} \tag{9.5.172}$$

Thus, finally

$$L_- Y_{l,m+1} = \frac{c_{l,m+1}}{c_{l,m}} Y_{l,m} = \sqrt{l(l+1) - m(m+1)}\,\hbar\, Y_{l,m} \, . \tag{9.5.173}$$

Now applying L_+ to this equation and using (9.5.126) we get

$$[l(l+1) - (m+1)^2 + (m+1)]\hbar^2\, Y_{l,m+1} = \sqrt{l(l+1) - m(m+1)}\,\hbar\, L_+ Y_{l,m} \tag{9.5.174}$$

or

$$L_+ Y_{l,m} = \sqrt{l(l+1) - m(m+1)}\,\hbar\, Y_{l,m+1} \, . \tag{9.5.175}$$

We can combine (9.5.173) and (9.5.175) into one very useful equation

$$L_\pm Y_{l,m} = \sqrt{l(l+1) - m(m\pm 1)}\,\hbar\, Y_{l,m\pm 1} \, . \tag{9.5.176}$$

Using L_- and iterating from $Y_{l,l}$ we finally get an explicit form for $Y_{l,m}$.

$$Y_{l,m} = \sqrt{\frac{(l+m)!}{(2l)!(l-m)!}}\,\hbar^{-(l-m)}\,L_-^{l-m}Y_{l,l}\ . \tag{9.5.177}$$

Furthermore, using the explicit form for P_l^m we get that

$$P_l^0(\cos\theta) = \sqrt{\frac{2l+1}{2}}\frac{1}{2^l l!}\left(\frac{d}{d\cos\theta}\right)^l(\cos^2\theta-1)^l\ . \tag{9.5.178}$$

This we now recognize as a Rodrigues formula for the Legendre polynomials $P_l(\cos\theta)$ which are normalized such that

$$P_l^0(\cos\theta) = \sqrt{\frac{2l+1}{2}}P_l(\cos\theta)\ . \tag{9.5.179}$$

Writing out equation (9.5.177) explicitly in terms of the $P_l^m(\cos\theta)$ equation (9.4.114) becomes

$$Y_{l,m}(\theta,\varphi) = \left(\frac{2l+1}{4\pi}\cdot\frac{(l-m)!}{(l+m)!}\right)^{1/2} P_l^m(\cos\theta)e^{im\varphi}\ . \tag{9.5.180}$$

The $P_l^m(\cos\theta)$ for $m > 0$ may be defined by what is called *Ferrer's formula*, namely

$$P_l^m(x) = \frac{(-1)^m}{2^l l!}(1-x^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l \tag{9.5.181}$$

and the corresponding functions for $-m$, $m > 0$ are then defined by

$$P_l^{-m}(x) = (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(x)\ . \tag{9.5.182}$$

From the normalization condition (9.5.171) we can now deduce the relationship:

$$\begin{aligned}\int_0^\pi P_l^m(\cos\theta)P_{l'}^{m'}(\cos\theta)\,\sin\theta\,d\theta &= \int_{-1}^1 P_l^m(x)P_{l'}^{m'}(x)\,dx\\ &= \frac{2}{2l+1}\cdot\frac{(l+m)!}{(l-m)!}\delta_{l,l'}\ .\end{aligned} \tag{9.5.183}$$

Now using (9.5.180) and (9.5.182) it follows that

$$Y_{l,-m}(\theta,\varphi) = (-1)^m Y_{l,m}^*(\theta,\varphi)\ . \tag{9.5.184}$$

For later reference we list the spherical harmonics for $l = 0,1,2,3$ and $m \geq 0$. The spherical harmonics for negative m can be deduced from (9.5.184). The phases are as given by (9.5.167) and (9.5.177).

The utility of spherical harmonics extends to all systems with spherical symmetry. We shall have many occasions to use them in subsequent chapters. For the present we first examine several more aspects of angular momentum.

Table 9.1: The Spherical Harmonics $Y_{l,m}(\theta,\varphi)$.

$\underline{l=0}$ $\quad Y_{0,0} = \frac{1}{\sqrt{4\pi}}$

$\underline{l=1}$ $\quad Y_{1,1} = -\sqrt{\frac{3}{8\pi}} \sin\theta \, e^{i\varphi}$

$Y_{1,0} = \sqrt{\frac{3}{4\pi}} \cos\theta$

$\underline{l=2}$ $\quad Y_{2,2} = \sqrt{\frac{15}{32\pi}} \sin^2\theta \, e^{2i\varphi}$

$Y_{2,1} = -\sqrt{\frac{15}{8\pi}} \sin\theta \, \cos\theta \, e^{i\varphi}$

$Y_{2,0} = \sqrt{\frac{3}{8\pi}} (2\cos^2\theta - 1)$

$\underline{l=3}$ $\quad Y_{3,3} = -\sqrt{\frac{35}{64\pi}} \sin^3\theta \, e^{3i\varphi}$

$Y_{3,2} = \sqrt{\frac{105}{32\pi}} \sin^2\theta \, \cos\theta \, e^{2i\varphi}$

$Y_{3,1} = -\sqrt{\frac{21}{64\pi}} \sin\theta (5\cos^2\theta - 1) e^{i\varphi}$

$Y_{3,0} = \sqrt{\frac{7}{16\pi}} (5\cos^3\theta - 3\cos\theta)$

9.6 Rotations and Rotational Invariance

A rotation of a system about a point O is a displacement of all points P of the system such that the distance between any two points remains unchanged. Rotations of coordinate systems are implemented by orthogonal matrices. Thus if the point $\mathbf{x}'$ is obtained from $\mathbf{x}$ by a rotation R about the origin of the coordinate system we write

$$\mathbf{x}' = R\mathbf{x} \ , \tag{9.6.185}$$

or in component form

$$x'_i = \sum_j R_{ij} x_j \ . \tag{9.6.186}$$

Since we need invariance of the distance from the origin we get as a condition

$$\sum_i x'_i x'_i = \sum_{ijk} R_{ij} x_j R_{ik} x_k = \sum_k x_k x_k \ . \tag{9.6.187}$$

This requires that

$$\sum_i R_{ij} R_{ik} = \delta_{jk} \tag{9.6.188}$$

and states that R, considered as a matrix, is *orthogonal*, namely

$$RR^t = R^t R = 1 \ . \tag{9.6.189}$$

Here R^t is the “transpose” of the matrix R. This equation immediately shows that

$$\det R \, \det R^t = (\det R)^2 = 1 \; . \tag{9.6.190}$$

Thus,

$$\det R = \pm 1 \; . \tag{9.6.191}$$

However, if we use the parity operator (section 4.13)

$$P\mathbf{x} = -\mathbf{x} \tag{9.6.192}$$

we see that

$$\det P = -1 \; . \tag{9.6.193}$$

And if $\det R = -1$, then $\det(PR) = \det P \det R = 1$. Conversely any R with determinant -1 can be written as

$$R = PR_p \tag{9.6.194}$$

where

$$\det R_p = +1 \; . \tag{9.6.195}$$

We call rotations with determinant $+1$, proper rotations (hence the subscript p). Those with determinant -1 are called improper and involve a parity transformation. Henceforth we consider only proper rotations and drop the subscript p.

Suppose we are in a coordinate system S with a quantum mechanical single particle state specified by a wavefunction ψ. Then if we rotate the coordinate system to a new system S' we get a new state ψ'. Since none of the physics has changed we require that all probabilities remain unchanged. So if ϕ is another state and ϕ' its image after rotation, we require that

$$|(\phi', \psi')|^2 = |(\phi, \psi)|^2 \; . \tag{9.6.196}$$

E. Wigner [2] has shown that (9.6.196) implies that ϕ' , ψ' can always be obtained from ϕ , ψ by either unitary or antiunitary transformations. The antiunitary transformations correspond to having time-reversal (section 4.6). We therefore consider unitary transformations. Thus,

$$\psi' = U_R \, \psi \tag{9.6.197}$$

where

$$U_R^\dagger \, U_R = U_R \, U_R^\dagger = 1 \; . \tag{9.6.198}$$

Clearly, if R_1 and R_2 are two rotations such that

$$R_1 R_2 = R_3 \tag{9.6.199}$$

then we need

$$U_{R_1R_2} = U_{R_1}U_{R_2} = U_{R_3} \quad . \tag{9.6.200}$$

This statement together with (9.6.198) says that the operators $\{U_R\}$ yield a *unitary representation of the rotation group.*

To proceed further we need to decide on how to specify a rotation. There are very many different methods available. The procedure we use is to specify an axis of rotation (unit vector) $\hat{n}$ and angle of rotation θ about this axis. Consider a rotation about the z-axis by an angle θ, i.e., $R_{z,\theta}$. Then,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad . \tag{9.6.201}$$

If we apply this transformation to the coordinates of a wavefunction $\psi(x,y,z)$ we get

$$U_{z,\theta}\psi(x,y,z) = \psi(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta, z) \quad . \tag{9.6.202}$$

Now let $\theta \to 0$, i.e., we consider an infinitesimal transformation. Then, to lowest order in θ

$$U_{z,\theta}\psi(x,y,z) = \psi(x + y\theta, -x\theta + y, z) \quad . \tag{9.6.203}$$

Taylor expanding (again to first order in θ) we obtain

$$\begin{aligned} U_{z,\theta}\psi(x,y,z) &= \psi(x,y,z,) + \theta\left[y\frac{\partial}{\partial y} - x\frac{\partial}{\partial x}\right]\psi(x,y,z) \\ &= \psi(x,y,z,) - \frac{i\theta}{\hbar}L_z\psi(x,y,z) \ . \end{aligned} \tag{9.6.204}$$

Thus, to lowest order in θ

$$U_{z,\theta} = 1 - \frac{i\theta}{\hbar}L_z \quad . \tag{9.6.205}$$

Similarly, we find for $\hat{n} = \hat{e}_x$ or $\hat{e}_y$ and infinitesimal θ that:

$$U_{x,\theta} = 1 - \frac{i\theta}{\hbar}L_x \ , \ U_{y,\theta} = 1 - \frac{i\theta}{\hbar}L_y \ . \tag{9.6.206}$$

Thus more generally for infinitesimal θ

$$U_{n,\theta} = 1 - \frac{i\theta}{\hbar}\hat{n}\cdot\mathbf{L} \ . \tag{9.6.207}$$

Now we use equations (9.6.200) and (9.6.207) for a finite angle θ together with an infinitesimal angle ϵ

$$\begin{aligned} U_{n,\theta+\epsilon} &= U_{n,\epsilon}U_{n,\theta} \\ &= \left(1 - \frac{i\epsilon}{\hbar}\hat{n}\cdot\mathbf{L}\right)U_{n,\theta} \quad . \end{aligned} \tag{9.6.208}$$

Thus, taking the limit as $\epsilon \to 0$ we get:

$$\lim_{\epsilon\to 0}\left[\frac{U_{n,\theta+\epsilon}-U_{n,\theta}}{\epsilon}\right]=\lim_{\epsilon\to 0}\left[-\frac{i}{\hbar}(\hat{n}\cdot\mathbf{L})U_{n,\theta}\right] . \tag{9.6.209}$$

Or since the limit defines a derivative we get

$$\frac{d}{d\theta}U_{n,\theta}=-\frac{i}{\hbar}(\hat{n}\cdot\mathbf{L})U_{n,\theta} , \tag{9.6.210}$$

where

$$U_{n,0}=1 . \tag{9.6.211}$$

Thus integrating we find

$$U_{n,\theta}=\exp\{-\frac{i}{\hbar}(\hat{n}\cdot\mathbf{L})\theta\} . \tag{9.6.212}$$

From (9.6.210) and (9.6.212) we see that the angular momentum operators $1/\hbar\,\mathbf{L}$ are the *generators of rotations* in the sense of Stone's Theorem (section 6.6). As always, the function e^A of an operator A is defined by its power series expansion.

Now consider a classical Hamiltonian H that remains invariant under rotations of the coordinate system. Typically such an H will be of the form

$$H=\frac{\mathbf{p}^2}{2m}+V(|\mathbf{r}|)=\frac{\mathbf{p}^2}{2m}+V(r) . \tag{9.6.213}$$

If the rotational invariance is to remain in the transition to quantum mechanics, we require that the Schrödinger equation also remain invariant under rotations. Thus, we need that

$$H\psi=E\psi \tag{9.6.214}$$

should imply that $U_R\psi$ is a solution whenever ψ is a solution. But

$$U_RH\psi=U_RHU_R^{-1}U_R\psi=EU_R\psi . \tag{9.6.215}$$

So for $U_R\psi$ to be a solution of (9.6.214) requires that

$$U_RHU_R^{-1}=H . \tag{9.6.216}$$

This is the condition that the quantum mechanical Hamiltonian be invariant under rotations. Multiplying equation (9.6.216) on the right by U_R and using infinitesimal rotations about the x, y and z-axis respectively we obtain (problem 9.11) the equations

$$[H,\mathbf{L}]=0 . \tag{9.6.217}$$

Equation (9.6.217) is equivalent to (9.6.216) and states necessary and sufficient conditions for rotational invariance of the Hamiltonian H.

9.7 Spin Angular Momentum

In finding the eigenvalues of L^2 and L_z by algebraic techniques we obtained, equation (9.5.146), that the eigenvalue l could be half-integral as well as integral. We now examine the special case of $l = 1/2$. In this case **L** cannot represent orbital angular momentum since by solving the differential equation for the eigenvalues of L_z equation (9.3.85) we found that the eigenvalues m, equation (9.3.89), had to be integers. The case of angular momentum $1/2\ \hbar$ represents a new intrinsic or internal quantum number called *spin*. It is as much a fundamental property of a particle as its charge or mass. If the total angular momentum which we now call spin s, is $s = 1/2\ \hbar$ then the z-component of spin s_z can have eigenvalues $m_s\hbar$ with

$$m_s = s, s-1, \ldots, -|s| = 1/2, -1/2\ . \tag{9.7.218}$$

Thus, there are only these two eigenvalues of m_s, namely ±1/2, possible. We call the corresponding eigenfunctions u_+ and u_- *spinors.*

$$s^2 u_\pm = (s_x^2 + s_y^2 + s_z^2) u_\pm = \frac{1}{2}(\frac{1}{2}+1)\hbar^2\, u_\pm \tag{9.7.219}$$

and

$$s_z u_\pm = \pm\frac{1}{2}\hbar\, u_\pm\ . \tag{9.7.220}$$

If we further write

$$s_\pm = s_x \pm i s_y \tag{9.7.221}$$

and use equations (9.5.176) for $l = 1/2$ we get:

$$s_+ u_+ = 0\ \ ,\ \ s_- u_- = 0 \tag{9.7.222}$$

$$s_+ u_- = \sqrt{\frac{1}{2}(\frac{1}{2}+1) + \frac{1}{2}(-\frac{1}{2}+1)}\hbar u_+ = \hbar u_+ \tag{9.7.223}$$

$$s_- u_+ = \sqrt{\frac{1}{2}(\frac{1}{2}+1) - \frac{1}{2}(\frac{1}{2}-1)}\hbar u_+ = \hbar u_-\ \ . \tag{9.7.224}$$

Evaluating the matrix elements of s_z and $s_\pm$ on the two-dimensional subspace spanned by $u_\pm$ the result can be written as 2 matrices. This allows us to write the spin operators as matrix operators.

$$\begin{aligned} s_z &= \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ s_+ &= \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ s_- &= \frac{\hbar}{2}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \end{aligned} \tag{9.7.225}$$

where the spinors $u_\pm$ are explicitly given by

$$u_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad u_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad . \tag{9.7.226}$$

But,

$$s_x = \frac{1}{2}(s_+ + s_-) = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{9.7.227}$$

and

$$s_y = \frac{i}{2}(s_- - s_+) = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{9.7.228}$$

and multiplying out and adding we find

$$s^2 = s_x^2 + s_y^2 + s_z^2 = \frac{3}{4}\hbar^2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad . \tag{9.7.229}$$

Thus, we have a simple matrix representation of the spin 1/2 operators and eigenfunctions. It is conventional to introduce the following three matrices

$$\begin{aligned} \sigma_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma_y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ \sigma_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{9.7.230}$$

called the *Pauli matrices.* In terms of these we can express the spin matrices $\mathbf{s}$ by

$$\mathbf{s} = \frac{\hbar}{2}\sigma \quad . \tag{9.7.231}$$

By explicit multiplication it is easy to verify that the Pauli matrices anticommute. This means

$$\sigma_x\sigma_y + \sigma_y\sigma_x = \sigma_x\sigma_z + \sigma_z\sigma_x = \sigma_y\sigma_z + \sigma_z\sigma_y = 0 \ . \tag{9.7.232}$$

The commutation relations among the Pauli matrices can also be obtained either by explicit multiplication or by using (9.7.230) and the commutation relations for s_x, s_y, s_z. In either case we get

$$\begin{aligned} \sigma_x\sigma_y - \sigma_y\sigma_x &= 2i\sigma_z \\ \sigma_y\sigma_z - \sigma_z\sigma_y &= 2i\sigma_x \\ \sigma_z\sigma_x - \sigma_x\sigma_z &= 2i\sigma_y \ . \end{aligned} \tag{9.7.233}$$

Combining these results with (9.7.232) we get

$$\begin{aligned} \sigma_x\sigma_y &= i\sigma_z \\ \sigma_y\sigma_z &= i\sigma_x \\ \sigma_z\sigma_x &= i\sigma_y \ . \end{aligned} \tag{9.7.234}$$

Table 9.2: Typical Magnetons and g-Factors

Particle	Magneton (erg/gauss)	g-Factor
electron	0.927×10^{-20}	-2.00
proton	0.505×10^{-23}	2×2.79
neutron	0.504×10^{-23}	-2×1.91

To understand the effect of spin, consider a free electron. Since an electron has spin 1/2, the energy is now 4-fold not just 2-fold degenerate. The eigenfunctions of the Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} \tag{9.7.235}$$

are

$$\psi_{\mathbf{k}\uparrow}(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad \mathbf{p} = \hbar\mathbf{k} \,, \quad s_z = \hbar/2 \tag{9.7.236}$$

$$\psi_{\mathbf{k}\downarrow}(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} , \quad \mathbf{p} = \hbar\mathbf{k} \,, \quad s_z = -\hbar/2 \tag{9.7.237}$$

$$\psi_{-\mathbf{k}\uparrow}(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{x}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad \mathbf{p} = -\hbar\mathbf{k} \,, \quad s_z = \hbar/2 \tag{9.7.238}$$

$$\psi_{-\mathbf{k}\downarrow}(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{x}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} , \quad \mathbf{p} = -\hbar\mathbf{k} \,, \quad s_z = -\hbar/2 \quad . \tag{9.7.239}$$

They are just the product of the momentum and spin eigenfunctions since all three spin operators $\mathbf{s}$ commute with H. In general, the Hamiltonian may have a spin-dependent part and then $\mathbf{s}$ does not commute with H. In that case we have to solve a pair of coupled differential equations to obtain the eigenfunctions and eigenvalues of H. This occurs whenever an external magnetic field is applied because a particle with spin carries a magnetic moment μ proportional to the spin, i.e.,

$$\mu = g\frac{\mu_B}{\hbar}\mathbf{s} \,. \tag{9.7.240}$$

Here g is the so-called gyromagnetic ratio (a pure number usually of the order of 1 in magnitude) and μ_B is the *magneton* of the particle. Thus,

$$\mu_B = \frac{|q|\hbar}{2mc} \tag{9.7.241}$$

where q is the charge, m is the mass of the particle involved, and c is the speed of light. Typical numbers are given in Table 9.2.

In the case of the electron, the quantity

$$\frac{e\hbar}{2mc} = 0.927 \times 10^{-20} \ \text{erg/gauss} \tag{9.7.242}$$

is known as one "Bohr magneton". The value of

$$\frac{e\hbar}{2M_p c} = 0.505 \times 10^{-23} \ \text{erg/gauss} \tag{9.7.243}$$

where M_p the mass of the proton is known as one "nuclear magneton". We examine the effect of the electronic magnetic moment in Chapter 16. For the present we leave further considerations of spin.

9.8 Problems

9.1 Use algebraic techniques to evaluate the following expectation values as a function of time for a simple harmonic oscillator state which at $t = 0$ is given by

$$\psi(0, x) = Au_0(x) + Bu_1(x) + Cu_3(x)$$

where $u_0(x)$ is the ground state, $u_1(x)$ is the first excited state and $u_3(x)$ is the third excited state of a S.H.O.:
$\langle H\rangle$, $\langle p^2/2m\rangle$, $\langle 1/2\, kx^2\rangle$, $\langle p\rangle$, $\langle x\rangle$, $(\Delta p)^2$, $(\Delta x)^2$.

9.2 a) Compute the 3×3 matrices $(L_j)_{m,m'}$ $j = x, y, z$ corresponding to $l = 1$.
b) Show that they satisfy the cyclic commutation relation

$$[L_x, L_y] = i\hbar\, L_z \quad \text{etc.}$$

c) Furthermore show that each matrix L_j satisfies the characteristic equation

$$\left(L_j^2 - \hbar^2 1\right) L_j = 0$$

d) Evaluate in closed form the expression for the matrix $\exp(iL_z\theta)$.
Hint: Write out the series and resum it after using the characteristic equation to simplify. Compare this result to a rotation matrix corresponding to a rotation through an angle θ about the z-axis for the quantities corresponding to

$$\left(\frac{x+iy}{\sqrt{2}} \, , \, z \, , \, \frac{x-iy}{\sqrt{2}}\right) .$$

9.3 Diagonalize the following Hamiltonian, that is, find a unitary operator U that brings the following Hamiltonian to diagonal form.

$$H = E\,a^\dagger a + V(a + a^\dagger)$$

where E and V are constants and

$$[a, a^\dagger] = \beta^2 \ ,$$

a positive constant.
Hint: Try to transform to operators

$$b = ua + v \ , \ b^\dagger = u^* a^\dagger + v^*$$

where u and v are complex numbers, and recall the simple harmonic oscillator.

9.4 Evaluate matrix elements of the form $(Y_{l,m}, zY_{l',m'})$. These occur in the evaluation of dipole radiation rates.
Hint: Use the following recurrence relation:

$$(2l+1)xP_l^m(x) = (l+m)P_{l-1}^m(x) + (l-m+1)P_{l+1}^m(x) \ .$$

9.5 Let A be an operator such that

$$[A, L_x] = [A, L_y] = 0 \ .$$

Calculate $[A, L^2]$.
Hint: First compute $[A, L_z]$.

9.6 Let a wavefunction be given by

$$\psi = A\left[\frac{x^2 - y^2 + 2ixy}{x^2 + y^2 + z^2} + \frac{3z}{\sqrt{x^2 + y^2 + z^2}} + 5\right]\exp(-\alpha\sqrt{x^2 + y^2 + z^2}).$$

Find the probability of obtaining any (l, m) value.

9.7 a) A measurement of the z-component of angular momentum is made on a particle in a state of total angular momentum 1 and x-component 1. What are the probabilities for obtaining the values $1, 0, -1$?
b) For a particle in a state of total angular momentum 1 and z-component 1 a measurement is made of the component of angular momentum along an axis lying in the $x-z$ plane and making an angle θ with the z-axis. What is the probability of getting the values $1, 0, -1$?

9.8 Compute the commutators $[\mathbf{x}, \mathbf{L}]$ and $[\mathbf{p}, \mathbf{L}]$ and compare the results with $[\mathbf{L}, \mathbf{L}]$. What does this suggest about the commutator $[\mathbf{A}, \mathbf{L}]$ where $\mathbf{A}$ is an arbitrary vector operator? Note that each of the commutators above constitutes nine commutators.

9.9 *Formally* one can derive the relation

$$[L_z, \varphi] = i\hbar$$

and deduce from it that

$$\Delta L_z \, \Delta\varphi \geq \hbar/2 \,.$$

Now $\Delta\varphi$ is of necessity $\leq 2\pi$, and in an eigenstate of L_z we have $\Delta L_z = 0$. This violates

$$\Delta L_z \, \Delta\varphi \geq \hbar/2 \,.$$

Explain this apparent paradox.
Hint: Examine the domain of L_z on which it is self-adjoint. A similar argument also holds for $[p, x] = i\hbar$ and a particle confined to a finite interval on the line. See also Carruthers & Nieto [9.2].

9.10 A system is in a state of angular momentum given by

$$\Psi = aY_{1,1} + bY_{1,0} + cY_{1,-1}$$

where

$$|a|^2 + |b|^2 + |c|^2 = 1 \;.$$

a) Compute the expectation value of L_x.
Hint: equation (9.5.176) may be useful.
b) Compute the expectation value of L^2.
c) What are possible values of the coefficients a, b, c in order that

$$L_x\Psi = \hbar\Psi \;?$$

It may again be useful to recall that

$$L_x = \frac{1}{2}(L_+ + L_-) \;.$$

9.11 Starting from equation (9.6.216),

$$U_R \, H \, U_R^{-1} = H$$

and using

$$U = \exp\{\frac{i\epsilon}{\hbar}\hat{\mathbf{n}} \cdot \mathbf{L}\}$$

with $|\epsilon| << 1$ obtain to first order in ϵ that

$$[H, \hat{\mathbf{n}} \cdot \mathbf{L}] = 0 \,.$$

Hence conclude that

$$[H, \mathbf{L}] = 0 \;.$$

9.12 Show that for any eigenstate of the simple harmonic oscillator

$$(\Delta x)^2 = \langle x^2 \rangle$$

$$(\Delta p)^2 = \langle p^2 \rangle$$

and also show for the state with quantum number n that

$$(\Delta x)(\Delta p) = (n + 1/2)\hbar \ .$$

9.13 Show that if you have two operators A, B such that

$$[A,[A,B]] = [B,[A,B]] = 0$$

then

$$e^{A+B} = e^A \, e^B \, e^{-1/2[A,B]} \ .$$

Hint: Consider the operator

$$f(x) = e^{xA} \, e^{xB}$$

and show that

$$\frac{df}{dx} = (A + B + [A,B]x) \ f(x) \ .$$

Integrate this equation and obtain the desired result. This result is a special case of the *Baker-Campbell-Hausdorff formula.*

9.14 Show that if

$$[A,B] = \lambda A$$

then

$$A \, e^B = e^\lambda \, e^B \, A \ .$$

This formula can be used, for example, if the operator B is proportional to the Hamiltonian $\hbar\omega(a^\dagger a + 1/2)$ and the operator A is proportional to either a or $a^\dagger$.

9.15 A particle is in a state described by the wave function

$$\psi = \frac{1}{\sqrt{6}}[Y_{l,l} + 2iY_{l,l-1} - Y_{l,l-2}] \ .$$

a) What are the most probable values that would be obtained in a single measurement of L^2 and L_z? What are the possible values of L_z and what are their corresponding probabilities?
b) Compute $(\Delta L_z)^2$ the uncertainty in L_z for this state.
c) If the state is now simply $Y_{l,m}$ find the uncertainty $(\Delta L_z)^2$ in L_z.

9.16 In Ehrenfest's theorem you proved that

$$\frac{d}{dt}\langle p\rangle = -\langle\frac{dV}{dx}\rangle = \langle F(x)\rangle \,.$$

Suppose that

$$F(x) = Ax^2 \ , \ \ A = \text{constant} \,.$$

Given that the system of interest is a simple harmonic oscillator and the wave function at $t = 0$ is

$$\Psi(0,x) = \frac{1}{2}[\sqrt{3}u_n(x) + iu_{n+1}(x)]$$

compute the difference between $F(\langle x\rangle)$ and $\langle F(x)\rangle$ as a function of time.

9.17 Compute Δx and Δp for a particle in the state $\psi_n(x)$ of a simple harmonic oscillator. Also verify the uncertainty relation.

9.18 Show that if $\vec{A}$ and $\vec{B}$ are two vector operators that commute with the Pauli matrices $\vec{\sigma}$ then the following equation holds

$$(\vec{\sigma}\cdot\vec{A})(\vec{\sigma}\cdot\vec{B}) = \vec{A}\cdot\vec{B} + i\vec{\sigma}\cdot(\vec{A}\times\vec{B}) \ \ .$$

9.19 Verify equation (9.5.163)

$$I_l = \frac{2l}{2l+1}I_{l-1}$$

starting from

$$I_l = \int_0^{\pi} \sin^{2l}\theta \, \sin\theta \, d\theta \ \ .$$

Bibliography

[9.1] C. van Winter, Orbital Angular Momentum and Group Representations - Annals of Phys. N.Y. **47**, 232-274 (1968).

[9.2] Several other papers which also consider the question of the single-valuedness of the wavefunction are listed below:
E. Merzbacher, Am. J. Phys. **30**, 237 (1962).
M.L. Whippmann, Am. J. Phys. **34**, 656 (1966).
W. Pauli, Helv. Phys. Acta **12**, 147 (1939).
P. Carruthers and M.M. Nieto - Phase and Angle Variables in Quantum Mechanics, Rev. Mod. Phys. **40**, 411 (1968).
E.T. Whittacker and G.N. Watson, *A Course of Modern Analysis*, Cambridge at he University Press, 4th edition (1963).
P.M. Morse and H. Feshbach, *Methods of Theoretical Physics* - McGraw-Hill Book Co., Inc. (1953).

Chapter 10

Central Force Problems

10.1 Introduction

There are many systems such that the potential is a function of only the distance r from the centre of force. In these cases the Hamiltonian is invariant under rotations and thus commutes with all components of the angular momentum operator. Such problems are called central force problems. We begin by extracting the angular variables (separation of variables) by using the eigenfunctions $Y_{l,m}$ of L^2 and L_z. The Schrödinger equation is thus effectively reduced to an equivalent Schrödinger equation in one variable. We then proceed to solve this radial Schrödinger equation for three different potentials: the infinite square well, the isotropic harmonic oscillator and the Coulomb potential. The isotropic simple harmonic oscillator is solved in Cartesian coordinates as well as in spherical coordinates. This latter solution allows us to introduce the parity operator in spherical coordinates. The associated Laguerre polynomials also arise naturally in these solutions. The hydrogenic atom (Coulomb potential) is discussed next. There we also give a further discussion of the associated Laguerre polynomials. Finally we show how to reduce a two-body problem to an equivalent one-body problem by extracting the centre of mass motion.

10.2 The Radial Equation

In this chapter we only consider Hamiltonians of the form

$$H = \frac{\mathbf{p}^2}{2M} + V(r) \ . \tag{10.2.1}$$

Since this Hamiltonian shows spherical symmetry we already saw (problem 9.12) that

$$[H, \mathbf{L}] = 0 \ . \tag{10.2.2}$$

So, H commutes with every component of $\mathbf{L}$ and hence

$$[H, L^2] = 0 \ . \tag{10.2.3}$$

Thus, it is possible to diagonalize H, L^2, and L_z simultaneously. As always we could have chosen L_x or L_y instead of L_z, but the choice of L_z is conventional. If we now consider the eigenvalue problem

$$H\,\psi_{E,l,m} = E\,\psi_{E,l,m} \tag{10.2.4}$$

then we can proceed by separation of variables or use the results of the previous chapter. We take the latter approach and set

$$\psi_{E,l,m} = R_{E,l,m}(r)\,Y_{l,m}(\theta,\varphi) \ . \tag{10.2.5}$$

This is possible because

$$\mathbf{p}^2\psi = -\hbar^2\nabla^2\psi = -\hbar^2\frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) + \frac{L^2}{r^2}\psi \ . \tag{10.2.6}$$

Thus (10.2.4) reads

$$-\frac{\hbar^2}{2M}\frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) + \frac{L^2}{2Mr^2}\psi + V(r)\psi = E\,\psi \ . \tag{10.2.7}$$

We have dropped the m dependence of R since (10.2.7) shows that R is independent of m. Now, using (10.2.5) this becomes

$$-\frac{\hbar^2}{2M}\frac{1}{r}\frac{d^2}{dr^2}(rR_{E,l}(r)) + \left[\frac{l(l+1)\hbar^2}{2MR^2} + V(r)\right]R_{E,l}(r) = E\,R_{E,l}(r) \ . \tag{10.2.8}$$

If we now call

$$u_{E,l}(r) = r\,R_{E,l}(r) \tag{10.2.9}$$

we get:

$$-\frac{\hbar^2}{2M}\frac{d^2u_{E,l}(r)}{dr^2} + \left[\frac{l(l+1)\hbar^2}{2Mr^2} + V(r)\right]u_{E,l}(r) = E\,u_{E,l}(r) \ . \tag{10.2.10}$$

This is identical in form to the one-dimensional Schrödinger equation. The effective potential is given here by

$$V_{eff}(r) = \frac{l(l+1)\hbar^2}{2Mr^2} + V(r) \ . \tag{10.2.11}$$

There is, however, one very major and significant difference here, namely $0 \le r < \infty$ and $U(0) = 0$ as can be seen from (10.2.9). Thus, to make this problem completely equivalent to a one-dimensional problem we need an effective potential (see figure 10.1)

$$V_{eff}(r) = \begin{cases} \frac{l(l+1)\hbar^2}{2Mr^2} + V(r) & r > 0 \\ \infty & r \le 0 \end{cases} \ . \tag{10.2.12}$$

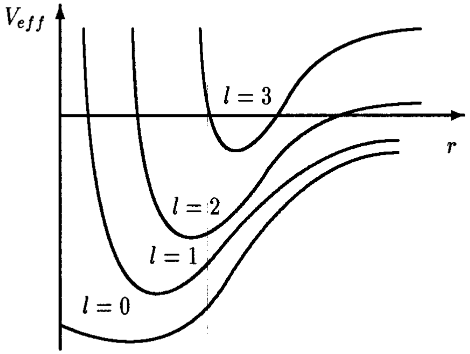

Figure 10.1: Typical effective potentials for a central force problem.

On the other hand, as we saw before (section 6.4), $\hbar/i\, d/dr$ is not a self-adjoint operator and has no self-adjoint extensions and thus cannot be considered as an observable, such as for instance "the momentum operator in the radial direction". A possible candidate for a radial momentum operator with the inner product

$$(f,g) = \int_0^\infty f^*(r)g(r)\, r^2\, dr \tag{10.2.13}$$

is

$$p_r f = \frac{\hbar}{i}\frac{1}{r}\frac{d}{dr}(rf) \ . \tag{10.2.14}$$

If we consider the deficiency indices of this operator for $\mathcal{L}_2(0,\infty)$ we get

$$p_r f = \pm i\hbar f \ . \tag{10.2.15}$$

Thus,

$$\frac{d}{dr}(rf) = \mp rf \ . \tag{10.2.16}$$

The solutions are

$$f_\pm = \frac{A}{r}\, e^{\mp r} \ . \tag{10.2.17}$$

Thus the deficiency indices are $(1,0)$ and p_r is not a possible observable. In fact there does not seem to be an observable corresponding to the radial momentum. Nevertheless p_r is a useful operator in that we can write the kinetic energy

$$T = \frac{\mathbf{p}^2}{2M} = \frac{p_r^2}{2M} + \frac{L^2}{2Mr^2} \ . \tag{10.2.18}$$

This is an easy way to remember the Laplace operator in spherical coordinates.

Let us now reconsider equation (10.2.10) with the effective potential V_{eff} given in (10.2.12). We are primarily interested in bound state problems so that $V(r)$ has to be attractive (negative). In fig.10.1 we have sketched $V_{eff}(r)$ for $r > 0$ for several different l values. Thus, as l increases, the depth of the well decreases and the well's minimum shifts to the right. Hence a particle tends to be less tightly bound. Also, since the wave function tends to concentrate above the minimum of the potential the most probable point for finding a particle moves further out. This is, of course, the same as in classical mechanics. A particle with high angular momentum is less tightly bound and tends to be in an orbit with a larger radius. We now consider some specific problems.

10.3 Infinite Square Well

The simplest problem we considered in one dimension was the infinite square well. We now reconsider the same problem in three dimensions. The potential we choose is

$$V = \begin{cases} 0 & r < a \\ \infty & r \geq a \end{cases} \quad . \tag{10.3.19}$$

Thus, we must solve the Schrödinger equation

$$\frac{\mathbf{p}^2}{2M}\psi = E\psi \tag{10.3.20}$$

with

$$\psi = 0 \quad \text{at} \quad r = a \; . \tag{10.3.21}$$

Since the potential is spherically symmetric we can set

$$\psi_{E,l,m} = R_{E,l}(r)Y_{l,m}(\theta,\varphi) \; . \tag{10.3.22}$$

Then,

$$L^2\psi_{E,l,m} = l(l+1)\hbar^2\,\psi_{E,l,m} \tag{10.3.23}$$

$$L_z\psi_{E,l,m} = m\hbar\,\psi_{E,l,m} \tag{10.3.24}$$

and

$$H\psi_{E,l,m} = E\psi_{E,l,m} \; . \tag{10.3.25}$$

This last equation reduces to an equation for $R_{E,l}(r)$, namely:

$$-\frac{\hbar^2}{2M}\frac{1}{r}\frac{d^2}{dr^2}(rR_{E,l}) + \left[\frac{l(l+1)\hbar^2}{2Mr^2} - E\right]R_{E,l} = 0 \; . \tag{10.3.26}$$

Setting,

$$k^2 = \frac{2ME}{\hbar^2} \tag{10.3.27}$$

and

$$x = kr \quad , \quad R_{E,l}(r) = y_{E,l}(x) \tag{10.3.28}$$

we get

$$\frac{1}{x}\frac{d^2}{dx^2}(xy_{E,l}) + \left[1 - \frac{l(l+1)}{x^2}\right] y_{E,l} = 0 \tag{10.3.29}$$

or

$$\left[\frac{d^2}{dx^2} + \frac{2}{x}\frac{d}{dx} + 1 - \frac{l(l+1)}{x^2}\right] y_{E,l} = 0 \ . \tag{10.3.30}$$

The solutions of this equation are the *spherical Bessel functions* which we examine in detail in section 19.6. They are written $j_l(x)$ and $n_l(x)$ where

$$j_l(x) = \sqrt{\frac{\pi}{2x}}\, J_{l+1/2}(x) \tag{10.3.31}$$

$$n_l(x) = (-1)^l \sqrt{\frac{\pi}{2x}}\, J_{-(l+1/2)}(x) \ . \tag{10.3.32}$$

Here $J_n(x)$ is the ordinary Bessel function which satisfies the differential equation

$$\left[\frac{d^2}{dx^2} + \frac{1}{x}\frac{d}{dx} + \left(1 - \frac{n^2}{x^2}\right)\right] J_n(x) = 0 \ . \tag{10.3.33}$$

To bring (10.3.30) to this form set

$$y = x^{-1/2} u \ . \tag{10.3.34}$$

Then u satisfies (10.3.33). Now the behaviours of $j_l(x)$ and $n_l(x)$ for small x are given by [1]

$$j_l(x) \to \frac{x^l}{(2l+1)!!} \quad \text{for} \quad x \to 0 \tag{10.3.35}$$

$$n_l(x) =\to -(2l-1)!!\frac{1}{x^{l+1}} \quad \text{for} \quad x \to 0 \ . \tag{10.3.36}$$

One of our boundary conditions is that $y_{E,l}(x)$ be finite at the origin. Thus, we have to drop the n_l solution and get

$$y_{E,l}(x) = A_l j_l(x) \ . \tag{10.3.37}$$

---

[1] All results pertaining to spherical Bessel functions are derived in section 19.6.

The equation for the energy eigenvalues is now given by the boundary condition

$$R_{E,l}(a) = 0 \tag{10.3.38}$$

or

$$j_l(a) = 0 \ . \tag{10.3.39}$$

Thus, if $\lambda_{n,l}$ is the nth zero of $j_l(x)$ then we have the energy$E_{n,l}$ given by

$$E_{n,l} = \frac{\hbar^2}{2Ma^2}\lambda_{n,l}^2 \ . \tag{10.3.40}$$

We shall not carry this problem any further. Instead we now consider the eigenvalue problem for the three dimensional simple harmonic oscillator.

10.4 Simple Harmonic Oscillator: Cartesian Coordinates

The Hamiltonian in this case is

$$H = \frac{\mathbf{p}^2}{2M} + \frac{1}{2}kr^2 \ . \tag{10.4.41}$$

We first solve this problem in Cartesian coordinates.

Clearly we can write H as

$$H = H_x + H_y + H_z \tag{10.4.42}$$

where

$$H_x = \frac{p_x^2}{2M} + \frac{1}{2}kx^2 \tag{10.4.43}$$

and similarly for H_y and H_z. Each of these is just a one-dimensional simple harmonic oscillator of the kind we solved before (section 9.2). Furthermore, all of these Hamiltonians are mutually commuting.

$$[H_x, H_y] = [H_x, H_z] = [H_z, H_y] = 0 \ . \tag{10.4.44}$$

Thus, we can diagonalize them simultaneously. Hence the eigenfunction ψ of

$$H\psi_n = E_n\psi_n \tag{10.4.45}$$

can be written

$$\psi = X(x)Y(y)Z(z) \tag{10.4.46}$$

where

$$H_x X_{n_1}(x) = E_{n_1} X_{n_1}(x) \tag{10.4.47}$$

and so forth. As we saw before, the eigenvalues are

$$\begin{aligned} E_{n_1} &= (n_1 + 1/2)\hbar\omega \\ E_{n_2} &= (n_2 + 1/2)\hbar\omega \\ E_{n_3} &= (n_3 + 1/2)\hbar\omega \ . \end{aligned} \tag{10.4.48}$$

Thus,

$$E_n \equiv E_{n_1,n_2,n_3} = (n_1 + n_2 + n_3 + 3/2)\hbar\omega \tag{10.4.49}$$

or

$$E_n = (n + 3/2)\hbar\omega \tag{10.4.50}$$

where

$$n = n_1 + n_2 + n_3 \ . \tag{10.4.51}$$

10.4.1 Degeneracy

We mentioned this concept before and simply recall the definition. An eigenvalue is *degenerate of order* g if there are g linearly independent eigenfunctions yielding this eigenvalue. In this case as a glance at (10.4.50) and (10.4.51) shows all eigenvalues except the ground-state (lowest state) are degenerate. Generally a degeneracy indicates that the Hamiltonian has some special symmetry property. In this case the Hamiltonian is spherically symmetric, hence the degeneracy. Another general property is that the ground-state is non-degenerate. We now examine the degeneracy of our oscillator more closely.
$n = 0$ In this case $n_1 = n_2 = n_3 = 0$ and this state is non-degenerate.
$n = 1$ In this case we have 3 possibilities: $n_1 = 1$, $n_2 = n_3 = 0$; $n_2 = 1$, $n_1 = n_3 = 0$; $n_3 = 1$, $n_1 = n_2 = 0$. Thus we have a three-fold degeneracy and the following eigenvalue solutions

$$\begin{aligned} H\psi_{1,0,0} &= 5/2\hbar\omega\psi_{1,0,0} \\ H\psi_{0,1,0} &= 5/2\hbar\omega\psi_{0,1,0} \\ H\psi_{0,0,1} &= 5/2\hbar\omega\psi_{0,0,1} \ . \end{aligned} \tag{10.4.52}$$

Lastly, $n = 2$. In this case the degeneracy is 6-fold.

The general formula for g, the order of the degeneracy, is

$$g = \frac{(n+2)(n+1)}{2} \ . \tag{10.4.53}$$

10.5 Simple Harmonic Oscillator: Spherical Coordinates

We now solve this problem again making explicit use of the spherical harmonics and orbital angular momentum. Since

$$H = \frac{\mathbf{p}^2}{2M} + \frac{1}{2}kr^2 \ , \tag{10.5.54}$$

we can write

$$\psi = \psi_{n,l,m} = R_{n,l}(r) Y_{l,m}(\theta, \varphi) \ . \tag{10.5.55}$$

This follows because

$$[L^2, H] = [L_z, H] = [L^2, L_z] = 0 \ . \tag{10.5.56}$$

In addition we find that the Hamiltonian (10.4.52) is invariant under parity transformations. Now under parity transformations both the position operator **r** and the momentum operator **p** change sign. Therefore under a parity transformation the angular momentum operator $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ also remains invariant. Thus,

$$P\,\mathbf{L} = \mathbf{L}\,P \ . \tag{10.5.57}$$

This allows us to choose simultaneous eigenstates of H, L^2, L_z, and P. Writing out the parity transformation

$$P\,\mathbf{r}\,P^{-1} = \mathbf{r}' = -\mathbf{r} \tag{10.5.58}$$

in spherical coordinates, we find (see also figure 10.2)

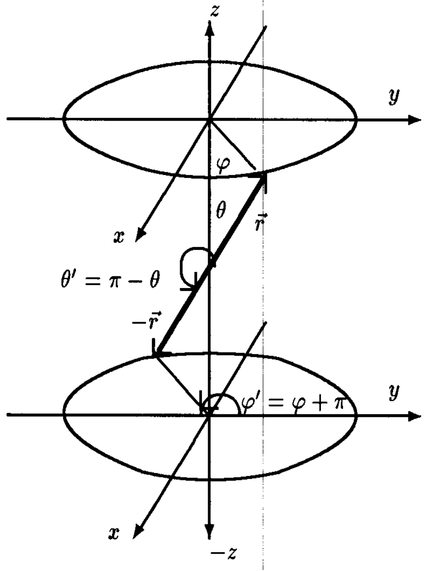

Figure 10.2: Parity transformation in spherical coordinates.

$$\begin{aligned} x' &= r' \sin\theta' \cos\varphi' = -r \sin\theta \cos\varphi \\ y' &= r' \sin\theta' \sin\varphi' = -r \sin\theta \sin\varphi \\ z' &= r' \cos\theta' = -r \cos\theta \ , \end{aligned} \tag{10.5.59}$$

so that

$$\begin{aligned} r' &= r \\ \theta' &= \pi - \theta \\ \varphi' &= \varphi + \pi \ . \end{aligned} \tag{10.5.60}$$

On the other hand, using the explicit form equations (9.5.180), (9.5.181) for the spherical harmonics, we find that

$$Y_{l,m}(\pi - \theta, \varphi + \pi) = (-1)^l \, Y_{l,m}(\theta, \varphi) \ . \tag{10.5.61}$$

Thus,

$$(P\,Y_{l,m})(\theta, \varphi) = (-1)^l \, Y_{l,m}(\theta, \varphi) \ . \tag{10.5.62}$$

This shows that the spherical harmonics are eigenstates of the parity operator with the eigenvalue +1 if l is even and −1 if l is odd.

Now we are ready to consider the radial equation for the simple harmonic oscillator. Setting, as usual,

$$rR_{n,l} = u_{n,l} \tag{10.5.63}$$

we find

$$\frac{d^2u}{dr^2} - \left[\frac{M^2\omega^2}{\hbar^2} r^2 + \frac{l(l+1)}{r^2}\right] u = -\frac{2ME}{\hbar^2} u \ . \tag{10.5.64}$$

To solve this equation we first consider the behaviour of u for $r \to 0$ and $r \to \infty$. These are respectively

$$u \sim r^{l+1} \tag{10.5.65}$$

and

$$u \sim \exp[-(M\omega)/(2\hbar) r^2] \, . \tag{10.5.66}$$

We again use Sommerfeld's polynomial method to find a solution of the form

$$u = A(r)\, r^{l+1} \, \exp[-(M\omega)/(2\hbar) r^2] \ . \tag{10.5.67}$$

Substituting this expression in equation (10.5.64), we get the following equation for $A(r)$

$$\frac{d^2A}{dr^2} + 2\left(\frac{l+1}{r} - \frac{M\omega}{\hbar} r\right) \frac{dA}{dr} + \left[\frac{2ME}{\hbar^2} - \frac{M\omega}{\hbar}(2l+3)\right] A = 0 \ . \tag{10.5.68}$$

Since we have already extracted the asymptotic behaviour for $u(r)$ the solutions for $A(r)$ must be a polynomial. Thus, we try

$$A = \sum_{n=0}^{M'} a_n \, r^n \ . \tag{10.5.69}$$

Substituting this expression into equation (10.5.68) and collecting terms with the same power of r we get

$$\sum_{n=0}^{M'} \left\{ n(n+l+2)a_{n+2} - \left[\frac{2M\omega}{\hbar}(n+l+3/2) - \frac{2ME}{\hbar^2} \right] a_n \right\} r^n = 0. \quad (10.5.70)$$

Thus, a solution exists if

$$a_{n+2} = \frac{\frac{2M\omega}{\hbar}(n+l+3/2) - \frac{2ME}{\hbar^2}}{n(n+l+2} a_n \; . \quad (10.5.71)$$

If furthermore,

$$a_{M'+2} = 0 \quad (10.5.72)$$

so that $a_{M'}$ is the highest term in the expansion (10.5.69) then the numerator on the right side of (10.5.71) must vanish for $n = M'$. This gives us the energy quantization

$$E_{M',l} = \hbar\omega(M' + l + 3/2) \quad M' = 0, 1, 2, \ldots \quad (10.5.73)$$

or

$$E_N = \hbar\omega(N + 3/2) \quad (10.5.74)$$

where

$$N = M' + l \; . \quad (10.5.75)$$

Thus, the energy is highly degenerate since we can have for a given M' all l values from 0 to $N - M'$. However, if we insist on parity as a good quantum number then according to equation (10.5.62) we can either have all even l or all odd l between 0 and $N - M'$ for a given M'. This means that M' must be an even integer, $M' = 2m$. It is now a simple matter to count the degeneracy and see that we get the same result as equation (10.4.53).

There is another way to proceed for the radial equation that allows us to make contact with functions (polynomials) that occur in the solution of the hydrogenic system considered next in section 10.5. The trick is to introduce a new independent variable

$$x = \frac{\hbar r^2}{M\omega} \quad \text{and} \quad y(x) = A(r) \; . \quad (10.5.76)$$

With these changes equation (10.5.68) becomes

$$x\frac{d^2y}{dx^2} + (l + 3/2 - x)\frac{dy}{dx} + \frac{1}{4}\left(\frac{2E}{\hbar\omega} - 2l - 3\right) y = 0 \; . \quad (10.5.77)$$

This is just a special case of the equation for the associated Laguerre polynomials $L^b_{a-b}(x)$, namely

$$x\frac{d^2}{dx^2}L^b_{a-b}(x) + (b + 1 - x)\frac{d}{dx}L^b_{a-b}(x) + (a - b)L^b_{a-b}(x) = 0 \; . \quad (10.5.78)$$

Thus to express $y(x)$ in terms of $L^b_{a-b}(x)$ we need the identification

$$b = l + 1/2 \tag{10.5.79}$$

$$a - b = \frac{1}{4}\left(\frac{2E}{\hbar\omega} - 2l - 3\right) \ . \tag{10.5.80}$$

It is a general property of equation (10.5.78) (see section 10.6.1) that for the solutions to be polynomials we need that $a - b = m$, an integer. Thus,

$$a = l + 1/2 + m \qquad m = 0, 1, 2, \ldots \tag{10.5.81}$$

and

$$\frac{1}{4}\left(\frac{2E}{\hbar\omega} - 2l - 3\right) = m \ . \tag{10.5.82}$$

So

$$E_{m,l} = \hbar\omega(l + 2m + 3/2) = \hbar\omega(N + 3/2) \tag{10.5.83}$$

where

$$N = l + 2m \ . \tag{10.5.84}$$

This is the same result we discovered above by using the parity operator, namely

$$E_N = \hbar\omega(N + 3/2) \ \ N = l + 2m \ \ m = 0, 1, 2, \ldots \ . \tag{10.5.85}$$

We introduced the associated Laguerre polynomials here in a rather ad hoc fashion. In the next section we encounter them again and discuss them in more detail.

10.6 The Hydrogenic Atom

We consider a particle of charge $-e$ in the electromagnetic field of a fixed charge Ze. Then

$$V(r) = -\frac{Ze^2}{r} \tag{10.6.86}$$

and

$$H = \frac{\mathbf{p}^2}{2M} - \frac{Ze^2}{r} \ . \tag{10.6.87}$$

Since the potential is central, angular momentum is conserved and we have

$$[L^2, H] = [L_z, H] = [L^2, L_z] = 0 \ . \tag{10.6.88}$$

Therefore, the Schrödinger equation

$$H\psi = E\psi \tag{10.6.89}$$

separates and we can write

$$\psi_{n,l,m} = R_{n,l}(r)\, Y_{l,m}(\theta, \varphi) \tag{10.6.90}$$

or putting

$$u_{n,l}(r) = rR_{n,l}(r) \tag{10.6.91}$$

as in equation (10.2.9) we get the radial equation

$$\left[-\frac{\hbar^2}{2M}\frac{d^2}{dr^2} + \frac{l(l+1)\hbar^2}{2Mr^2} - \frac{Ze^2}{r}\right] u = Eu \ . \tag{10.6.92}$$

It is convenient to introduce dimensionless quantities. To this end we define

$$\kappa^2 = -\frac{2ME}{\hbar^2} \ . \tag{10.6.93}$$

The minus sign is due to the fact that we are looking for bound states so that the energy E is negative. Now let

$$x = 2\kappa r \qquad u(r) = y(x) \ . \tag{10.6.94}$$

Then, with

$$\nu = \frac{MZe^2}{\kappa\hbar^2} \tag{10.6.95}$$

equation (10.6.92) becomes

$$y'' - \left[\frac{l(l+1)}{x^2} - \frac{\nu}{x} + \frac{1}{4}\right] y = 0 \ . \tag{10.6.96}$$

We again consider first the asymptotic forms of y and then use Sommerfeld's polynomial method. For large x (10.6.96) becomes:

$$y'' - \frac{1}{4}y \approx 0 \quad x \to \infty \,. \tag{10.6.97}$$

So,

$$y \sim e^{\pm x/2} \quad x \to \infty \,. \tag{10.6.98}$$

Since y must be square-integrable only the minus sign is acceptable. For small x the equation becomes

$$y'' - \frac{l(l+1)}{x^2} y \approx 0 \,. \tag{10.6.99}$$

The solutions are

$$y \sim x^{l+1} \quad \text{and} \quad y \sim x^{-l} \ . \tag{10.6.100}$$

Since we must have $y \to 0$ as $x \to 0$ only $y \sim x^{l+1}$ is acceptable. We therefore set

$$y = x^{l+1}\, e^{-x/2}\, v(x) \tag{10.6.101}$$

and substitute this into (10.6.96) to obtain an equation for $v(x)$. The resultant equation is

$$xv'' + (2l + 2 - x)v' + (\nu - l - 1)v = 0 \ . \tag{10.6.102}$$

This is precisely of the same form as equation (10.5.78) if we set

$$b = 2l + 1 \quad , \quad a = l + \nu \ . \tag{10.6.103}$$

To get polynomial solutions requires that

$$a - b = \nu - l - 1 \tag{10.6.104}$$

be a non-negative integer. Thus,

$$\nu = l + N = n \quad , \quad N = 1, 2, \dots \ . \tag{10.6.105}$$

This is precisely the quantization condition for the energy since

$$\nu = \frac{MZe^2}{\kappa\hbar^2} = \frac{MZe^2}{\hbar^2}\sqrt{-\frac{\hbar^2}{2ME}} = n \ . \tag{10.6.106}$$

Thus

$$E_n = -\frac{1}{2}M\frac{Z^2e^4}{\hbar^2}\frac{1}{n^2} \tag{10.6.107}$$

or introducing a dimensionless quantity called the *fine structure constant*

$$\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137} \tag{10.6.108}$$

we get

$$E_n = -\frac{1}{2}Mc^2Z^2\alpha^2\frac{1}{n^2} \quad n = 1, 2, \dots \ . \tag{10.6.109}$$

This is the same result we obtained using Bohr-Sommerfeld quantization (section 2.4). This equation may also be written

$$E_n = -\frac{Z^2e^2}{2a_0}\frac{1}{n^2} \tag{10.6.110}$$

where

$$a_0 = \frac{\hbar^2}{Me^2} \tag{10.6.111}$$

is the *Bohr radius.*

10.6.1 Laguerre Polynomials

We now return to a systematic study of equation (10.5.78), or equivalently (10.6.102). If we try a series solution

$$L^b_{a-b}(x) = 1 + a_1 x + a_2 x^2 + \ldots + a_{N-1} x^{N-1} \tag{10.6.112}$$

and substitute this into equation (10.5.78) and equate the coefficients of equal powers of x we get:

$$(n+1)(n+1+b)a_{n+1} = (n+b-a)a_n \ . \tag{10.6.113}$$

If we want $a_{N-1} \neq 0$ but $a_N = 0$ then we need that

$$N - 1 + b - a = 0 \tag{10.6.114}$$

or

$$N = a - b + 1 \ . \tag{10.6.115}$$

This is the condition we used in the previous section to ensure that our solutions are polynomials. If our solutions were not polynomials we would have to examine the convergence of the series $\sum a_n x^n$, where equation (10.6.113) yields a recursion relation for the coefficients a_n. From (10.6.113) we see that

$$\frac{a_n}{a_{n-1}} = \frac{n+b-a-1}{n(n+b)} \to \frac{1}{n} \quad \text{as} \quad n \to \infty \ . \tag{10.6.116}$$

Thus, the asymptotic behaviour of the series would be such that

$$\frac{a_n}{a_{n-1}} = \frac{1}{n} \tag{10.6.117}$$

or

$$a_n = \frac{1}{n!} \ . \tag{10.6.118}$$

This would lead to behaviour like e^x and in view of the asymptotic behaviour displayed in equation (10.6.101) is unacceptable. Now consider equation (10.5.78) with $y = L^b_{a-b}(x)$.

$$x\frac{d^2y}{dx^2} + (b+1-x)\frac{dy}{dx} + (a-b)y = 0 \ . \tag{10.6.119}$$

We look for a convenient integral representation for the solution of this equation. Thus, we try

$$y(x) = \oint_C e^{-xt} f(t)\, dt \tag{10.6.120}$$

where we shall later choose the integration path for our convenience. Substituting into (10.5.78) and differentiating under the integral sign yields:

$$\oint_C f(t)\ \left[xt^2 - (b+1-x)t + a - b\right] e^{-xt}\, dt = 0 \ . \tag{10.6.121}$$

This can be rewritten as

$$\oint_C f(t)\left[-t(t+1)\frac{d}{dt}+a-b-(b+1)t\right]e^{-xt}\,dt=0\ . \tag{10.6.122}$$

We now rewrite the first term as a total differential by subtracting and adding a term $-e^{-xt}d/dt[f(t)t(t+1)]$. Then we get

$$\begin{aligned}&\oint_C \frac{d}{dt}\left[-e^{-xt}f(t)t(t+1)\right]\,dt+\\&\oint_C e^{-xt}\left\{t(t+1)\frac{df}{dt}+[a-b+1-(b-1)t]f(t)\right\}\,dt=0\ .\end{aligned} \tag{10.6.123}$$

If the contour C forms a closed path then the first integral vanishes. To make the second integral vanish we now simply choose f to satisfy

$$t(t+1)\frac{df}{dt}=-[a-b+1-(b-1)t]f(t)\ . \tag{10.6.124}$$

The solution to this equation is:

$$f(t)=A\frac{(t+1)^a}{t^{a-b+1}}\ . \tag{10.6.125}$$

Now in order to be able to close the contour we must not cross any branch points. Thus we need $a=$ integer, $a-b+1=$ integer. This allows us to choose the contour to be a circle enclosing the origin. We then find

$$y(x)=A\oint e^{-xt}\frac{(t+1)^a}{t^{a-b+1}}\,dt=0\ . \tag{10.6.126}$$

Now we have only the contribution from the simple pole at $t=0$. This residue is obtained by expanding the exponential and binomial

$$e^{-xt}=\sum_{n=0}^{\infty}\frac{(-x)^n}{n!}t^n \tag{10.6.127}$$

and

$$(t+1)^a=\sum_{r=0}^{a}\begin{pmatrix}a\\r\end{pmatrix}t^r \tag{10.6.128}$$

so that the integrand is

$$\sum_{n=0}^{\infty}\sum_{r=0}^{a}\frac{(-x)^n}{n!}\begin{pmatrix}a\\r\end{pmatrix}t^{n+r+b-a-1}\ . \tag{10.6.129}$$

Thus we have a simple pole whenever

$$n+r+b-a=0 \tag{10.6.130}$$

or

$$n=a-b-r\ . \tag{10.6.131}$$

The contribution of this pole is

$$\oint e^{-xt}\frac{(t+1)^a}{t^{a-b+1}}\,dt = 2\pi i\frac{(-x)^{a-b-r}}{(a-b-r)!}\begin{pmatrix} a \\ r \end{pmatrix} \ . \tag{10.6.132}$$

Thus, up to normalization, we now have

$$L^b_{a-b} = A\sum_{r=0}^{a-b} 2\pi i\frac{(-x)^{a-b-r}}{(a-b-r)!}\begin{pmatrix} a \\ r \end{pmatrix} \ . \tag{10.6.133}$$

The normalization is by convention such that

$$L^0_a(0) \equiv L_a(0) = a! \ . \tag{10.6.134}$$

The functions $L_a(x)$ are known as the *Laguerre polynomials*. Substituting this into (10.6.133) we get

$$A = \frac{a!}{2\pi i} \ . \tag{10.6.135}$$

So we have

$$L_a(x) \equiv L^0_a(x) = \frac{a!}{2\pi i}\oint e^{-xt}\,\frac{(t+1)^a}{t^{a+1}}\,dt \ . \tag{10.6.136}$$

We can rewrite this as

$$\begin{aligned} L_a(x) &= \frac{a!}{2\pi i}e^x(-1)^a\frac{d^a}{dx^a}\oint\frac{e^{-x(t+1)}}{t^{a+1}}\,dt \\ &= (-1)^a\frac{a!}{2\pi i}e^x\frac{d^a}{dx^a}\oint\frac{e^{-xt}}{t^{a+1}}\,dt\,e^{-x} \end{aligned} \tag{10.6.137}$$

or, evaluating the integral,

$$L_a(x) = e^x\frac{d^a}{dx^a}\left(x^a e^{-x}\right) \ . \tag{10.6.138}$$

This is a Rodrigues' formula for the Laguerre polynomials $L_a(x)$.

To get the associated Laguerre polynomials $L^b_{a-b}(x)$ we first note that the Laguerre polynomials $L_a(x)$ satisfy the equation

$$\left[x\frac{d^2}{dx^2} + (1-x)\frac{d}{dx} + a\right]L_a(x) = 0 \ . \tag{10.6.139}$$

If we differentiate this equation b times we find that

$$y^b_a = \frac{d^bL_a}{dx^b} \tag{10.6.140}$$

satisfies

$$\left[x\frac{d^2}{dx^2} + (b+1-x)\frac{d}{dx} + (a-b)\right]y^b_a(x) = 0 \tag{10.6.141}$$

which is the same as equation (10.5.78). Thus, we have

$$L^b_{a-b}(x) = c^b_a \frac{d^b L_a}{dx^b} \tag{10.6.142}$$

where c^b_a is a constant which by convention is chosen as $(-1)^b$. Thus,

$$L^b_{a-b}(x) = (-1)^b \frac{d^b L_a}{dx^b} \tag{10.6.143}$$

or using (10.6.126) and (10.6.135)

$$L^b_{a-b}(x) = \frac{a!}{2\pi i} \oint e^{-xt} \frac{(t+1)^a}{t^{a-b+1}}\, dt \ . \tag{10.6.144}$$

In terms of these functions we are now ready to write down the bound-state wave functions for the hydrogenic atom.

A note of caution is in order.

The functions that we have written as $L^b_{a-b}(x)$ are also written sometimes as $L^b_a(x)$. Since both notations are fairly common, care must be exercised when mixing formulas from different books. For reference we list some of the more common textbooks and their notation. The notation used in this book, $L^b_{a-b}(x)$, is the same as that used by Liboff, Merzbacher, and Messiah [10.4]. The notation $L^b_a(x)$ is used by: Pauling and Wilson, Schiff, and Tomonaga [10.5]. The notation used by Gradshteyn and Ryzhik [10.6] is related to ours as follows: $L^q_p(x)$ (Capri) $= (p+q)!L^q_p(x)$ (Gradshteyn and Ryzhik).

Up to normalization the hydrogen atom wave functions are given by (our notation)

$$\psi_{n,l,m}(r,\theta,\varphi) = A_{n,l}\, e^{-x/2}\, x^l\, L^{2l+1}_{n-l-1}(x)\, Y_{l,m}(\theta,\varphi) \tag{10.6.145}$$

where we have kept $x = 2\kappa r$. The constant $A_{n,l}$ is the normalization constant and may be evaluated in a number of ways (most often a generating function is used). The result is

$$\begin{aligned} \psi_{n,l,m}(\mathbf{r}) \ &= \ \sqrt{\left(\frac{2Z}{na_0}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}} \left(\frac{2Zr}{na_0}\right)^l \\ &\times \ e^{-Zr/na_0}\, L^{2l+1}_{n-l-1}\left(\frac{2Zr}{na_0}\right) Y_{l,m}(\theta,\varphi) \end{aligned} \tag{10.6.146}$$

where, as before, a_0 is the Bohr radius.

$$a_0 = \frac{\hbar^2}{Me^2} \ . \tag{10.6.147}$$

Since the hydrogen atom has a high degree of symmetry (it is invariant under rotations) we expect at least the usual $(2l+1)$-fold degeneracy associated with a central potential. In fact the degeneracy is even higher. This so-called "accidental" degeneracy is due to the fact that the hydrogen Hamiltonian does not change under an even larger group of transformations than just the three-dimensional rotations.

We now calculate g_n, the degree of the degeneracy for each level n. Since the value of E_n depends only on n we have degeneracy with respect to both m (rotational) and l (accidental). For each fixed value of n, l can vary from 0 to $n-1$ and for each of these values m can vary over $2l+1$ values from $-l$ to l. Thus, the degree of degeneracy (ignoring spin) is given by:

$$g_n = \sum_{l=0}^{n-1}(2l+1) = 2\frac{n(n-1)}{2} + n = n^2 \ . \tag{10.6.148}$$

The problem we have solved yields only the bound state (square-integrable) wavefunctions for the hydrogenic atom. These do *not* form a complete set. In addition there are solutions that are only δ-function normalizable. These solutions correspond to continuous positive values of the energy and represent solutions for particles scattered by a Coulomb potential. Due to the fact that the Coulomb potential decreases only very slowly (as r^{-1}), these solutions are rather complicated and will not be considered by us. They are expressible in terms of hypergeometric functions but are usually avoided. In practice the Coulomb potential is usually "screened" by other charges and thus yields an effective potential that decreases more rapidly. This yields a simpler scattering problem and is considered in section 19.5 where we obtain the quantum mechanical analogue of Rutherford scattering by this means. In our treatment of the hydrogen atom we have not yet justified the use of a fixed center of force. We now do so and show how this is accomplished for a general two-body problem. The procedure is identical to that used in classical mechanics for removing the center of mass motion.

10.7 Reduction of the Two-Body Problem

Consider a general Hamiltonian for two particles interacting only with each other. Then, using the fact that space is homogeneous (no preferred origin) restricts the possible interaction potential to a function of $\mathbf{r}_1 - \mathbf{r}_2$ where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the position vectors of particles 1 and 2 respectively. The Hamiltonian is therefore of the form

$$H = \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} + V(\mathbf{r}_1 - \mathbf{r}_2) \ . \tag{10.7.149}$$

We now introduce the centre of mass and relative coordinates

$$\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} \tag{10.7.150}$$

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \ . \tag{10.7.151}$$

It is also convenient to introduce the total and the reduced mass

$$M = m_1 + m_2 \tag{10.7.152}$$

$$m = \frac{m_1 m_2}{m_1 + m_2} \tag{10.7.153}$$

as well as the total momentum

$$\mathbf{P} = -i\hbar\nabla_R \tag{10.7.154}$$

and the relative momentum

$$\mathbf{p} = -i\hbar\nabla_r \ . \tag{10.7.155}$$

The Hamiltonian can then be rewritten (problem 10.3) to read

$$H = \frac{\mathbf{P}^2}{2M} + \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) \ . \tag{10.7.156}$$

Furthermore

$$[\mathbf{P}, \mathbf{p}] = 0 \tag{10.7.157}$$

and

$$[\mathbf{P}, H] = 0 \ . \tag{10.7.158}$$

Thus, we can diagonalize $\mathbf{P}$ and H simultaneously. This amounts, of course, to extracting the centre of mass motion. In fact, if we call

$$H_0 = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) \ . \tag{10.7.159}$$

Then H_0, H and $\mathbf{P}$ can all be simultaneously diagonalized to give

$$\mathbf{P}\,\phi(\mathbf{R}) = \hbar\mathbf{K}\,\phi(\mathbf{R}) \tag{10.7.160}$$

$$H_0\psi(\mathbf{r}) = E_0\psi(\mathbf{r}) \tag{10.7.161}$$

and

$$H_0\Psi(\mathbf{R},\mathbf{r}) = \left(\frac{\hbar^2\mathbf{K}^2}{2M} + E_0\right)\Psi(\mathbf{R},\mathbf{r}) \ . \tag{10.7.162}$$

This is, in fact, accomplished simply by separation of variables, i.e.

$$\Psi(\mathbf{R},\mathbf{r}) = \phi(\mathbf{R})\psi(\mathbf{r}) \ . \tag{10.7.163}$$

Then

$$\phi(\mathbf{R}) = \frac{1}{(2\pi)^{3/2}}\, e^{i\mathbf{K}\cdot\mathbf{R}} \ . \tag{10.7.164}$$

Equation (10.7.161) is now nothing more than the Schrödinger equation for a particle in a fixed centre of force. Thus, we get the equivalent one-body problem from a given two-body problem by simply using the reduced mass in the relative coordinate system. For the hydrogen atom this amounts to replacing the electron mass m_e by the *reduced mass*

$$m = \frac{m_e m_p}{m_e + m_p} \tag{10.7.165}$$

where m_p, the proton mass is about 1840 m_e. Thus, this produces a correction of about 0.05% and is well within the limits of observable effects.

This concludes our treatment of exactly solvable problems. There are several more classes of potentials for which closed-form solutions are known. The principal merit of these exact solutions is that they produce a point of departure for approximate solutions. This will be the subject matter of a later chapter. In preparation for this we next develop some more formalism and in the process improve our notation as well.

10.8 Problems

10.1 Solve the isotropic simple harmonic oscillator problem in two dimensions in both Cartesian and cylindrical coordinates.
Hint: L_z commutes with the Hamiltonian.

10.2 Consider the attractive potential

$$V(r) = -V_0\, e^{-\alpha r}$$

for $l = 0$. This is one of the few solvable problems.
Hint: Change variables to $u = e^{-\alpha r}$. The resultant equation is Bessel's equation. Discuss carefully the boundary conditions to be obeyed by $\phi(u) = R(r)$.

10.3 Show that the Hamiltonian given in (10.7.149) reduces to the Hamiltonian in (10.7.156) under the transformations (10.7.150) - (10.7.153).

10.4 A particle is in a spherical potential well

$$V(r) = \begin{cases} -V_0 & \text{for} \quad r < a \\ 0 & \text{for} \quad r > a \end{cases} .$$

Find the transcendental equation which yields the energy eigenvalue for the state with angular momentum l. What is the minimum degeneracy of this state? If a proton and neutron are bound in an $l = 0$ state with an energy of 2.2 MeV, determine V_0 given that $a \approx 2 \times 10^{-13}$ cm.

10.5 A particle is free to move on the surface of a circular cylinder of radius R. The Laplacian in cylindrical coordinates is given by:

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial z^2} .$$

Find the energy eigenvalues and eigenfunctions for this motion.

10.6 An electron in the Coulomb field of a proton is in a state described by the ket:

$$\frac{1}{6}\left[4|1,0,0\rangle + 3|2,1,1\rangle - |2,1,0\rangle + \sqrt{10}|2,1,-1\rangle\right]$$

where the labelling is $|n, l, m\rangle$. Find
a) The expectation value of the energy.
b) The expectation value of L^2.
c) The expectation value of L_z.

10.7 Show that under a parity transformation

$$\theta \to \theta' = \pi - \theta$$

$$\varphi \to \varphi' = \varphi + \pi \quad .$$

Bibliography

[10.1] The original papers treating quantization and the hydrogen atom are still very readable. They are:
E. Schrödinger, Annalen der Physik **79** 361, 489, (1926).
ibid **81**, 109, (1926).
Many detailed solutions of central force problems are to be found in *Pauli Lectures on Physics Vol.5 Wave Mechanics* edited by C.P. Enz, The MIT Press, Cambridge, Mass. 1973.

[10.2] For some of the difficulties and interpretation of the "radial momentum operator" see:
F.S. Crawford Jr., Amer. J. Phys. **32**, 611, (1964).

[10.3] The classical limit of the hydrogen atom is discussed in:
L.S. Brown, Amer. J. Phys. **41**, 525, (1973).

[10.4] Further references are:
R.L. Liboff, *Introduction to Quantum Mechanics*, Holden-Day Inc. (1980).
E. Merzbacher, *Quantum Mechanics* 2nd edition, John Wiley and Sons, Inc. (1970).
A. Messiah, *Quantum Mechanics*, John Wiley and Sons, Inc. (1966).

[10.5] L. Pauling and E.B. Wilson, *Introduction to Quantum Mechanics*, McGraw-Hill, N.Y. (1935).
L. Schiff, *Quantum Mechanics* 3rd edition, McGraw-Hill, N.Y. (1968).
S. Tomonaga, *Quantum Mechanics*, North-Holland (1966).

[10.6] I.S. Gradshteyn and I.M. Ryzhik, *Tables of Integrals, Series and Products*, Academic Press (1965).

Chapter 11

Transformation Theory

11.1 Introduction

The choice of a set of coordinate axes or basis is completely arbitrary in Euclidean space. The same thing is true in Hilbert space. It is therefore very useful to know how to change from one basis to another. This is known as "transformation theory". Since the basis set is completely arbitrary it is also useful to work as much as possible in a basis independent manner. This is what we do in Euclidean space when we work with equations for vectors rather than with equations for the components of vectors. There is a completely analogous procedure available in Hilbert space involving a special vector notation called Dirac notation. This powerful notation is discussed in this chapter. Transformations from one basis set to another are carried out by means of unitary transformations. If we also permit time-dependent unitary transformations we can throw all, or part, of the time dependence onto the operators. We refer to these time transformed representations as "pictures" and discuss three of them, the Schrödinger, Heisenberg and Dirac in detail.

11.2 Rotations in a Vector Space

Since by definition, a *basis set* is any *complete orthonormal set*, say $\{u_n(x_1, x_2, x_3)\}$, we can expand any given wave function $\psi(x_1, x_2, x_3)$ in terms of such a basis set. The index n may be a multi-index consisting of a set of discrete indices as in the three-dimensional oscillator where $n = (n_1, n_2, n_3)$ or it may be a set of continuous indices or a mixture of discrete and continuous indices. We shall use a summation sign as a generic symbol for summation over discrete and integration over continuous indices. Thus, in the case stated,

$$\psi(x_1, x_2, x_3) = \sum_n a_n u_n(x_1, x_2, x_3) \tag{11.2.1}$$

where

$$a_n = (u_n, \psi) \ . \tag{11.2.2}$$

It is clear that once we have picked the basis set $\{u_n\}$, the wave function is completely specified by the set of ordered numbers $(a_1, a_2, \ldots)$. This is completely analogous to the representation of vectors by ordered n-tuples in ordinary analytic geometry. For example in $\mathcal{E}_3$ we may pick an orthonormal triad $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ Then we can write any vector $\mathbf{v}$ as

$$\mathbf{v} = \sum_{n=1}^{3} a_n \hat{e}_n \tag{11.2.3}$$

where

$$a_n = (\hat{e}_n, \mathbf{v}) \ . \tag{11.2.4}$$

It is now quite common to suppress the basis vectors $\{\hat{e}_n\}$ and write

$$\mathbf{v} = (a_1, a_2, a_3) \ . \tag{11.2.5}$$

The change from one basis set to another is accomplished in $\mathcal{E}_3$ by means of rotations or orthogonal transformations. To see this consider a second basis set $\{\hat{f}_1, \hat{f}_2, \hat{f}_3\}$ obtained from the set $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ by a rotation R. Thus, we have:

$$\hat{f}_n = \sum_{l=1}^{3} R_{nl} \hat{e}_l \ . \tag{11.2.6}$$

The statement, that the set $\{\hat{f} f_n\}$ is still orthonormal, when written out, reads

$$\begin{aligned} (\hat{f}_n, \hat{f}_m) &= \delta_{mn} \\ &= \left(\sum_k R_{nk} \hat{e}_k, \sum_l R_{ml} \hat{e}_l \right) \\ &= \sum_{k,l} R_{nk} R_{ml} (\hat{e}_k, \hat{e}_l) \\ &= \sum_{k,l} R_{nk} R_{ml} \delta_{kl} \\ &= \sum_k R_{nk} R_{mk} \ . \end{aligned} \tag{11.2.7}$$

In terms of the matrix R, equation (11.2.7) reads:

$$RR^t = R^t R = 1 \ . \tag{11.2.8}$$

Here the superscript t means "transpose". Equation (11.2.8) states that the rotation matrices R are orthogonal. Conversely if the matrices R are orthogonal so that equation (11.2.8) holds then the vectors $\hat{f}_n$ defined by equation (11.2.6) also form an orthonormal basis set if the $\hat{e}_n$ do. A completely analogous condition holds for vectors in Hilbert space $\mathcal{H}$. In this case, since the vectors are complex, we have complex orthogonal or unitary transformations. To see how this works consider two different orthonormal basis sets $\{u_n\}$ and $\{v_n\}$. Then, since the sets are by definition complete we can expand one set in terms of the other. Thus, we have

$$v_n = \sum_m V_{nm} u_m \tag{11.2.9}$$

and

$$u_n = \sum_m V^{-1}_{nm} v_m \quad . \tag{11.2.10}$$

That V^{-1} exists is obvious from the fact that the u_n may be expanded in terms of the v_n. Now using the orthonormality of the sets we get

$$\begin{aligned}
(v_n, v_m) &= \delta_{mn} \\
&= \sum_{k,l} V^*_{nk} V_{ml} (u_k, u_l) \\
&= \sum_{k,l} V^*_{nk} V_{ml} \delta_{kl} \\
&= \sum_k V^*_{nk} V_{mk}
\end{aligned} \tag{11.2.11}$$

or in operator form

$$V\, V^{*t} \equiv V V^{\dagger} = 1\ . \tag{11.2.12}$$

In a similar fashion using (11.2.10) and expanding u_n in terms of the v_n we obtain:

$$\delta_{nm} = \sum_k V^{-1}_{nk} V^{*\,1}_{mk} \tag{11.2.13}$$

or in operator form

$$V^{-1}\, V^{\dagger\, -1} = 1\ . \tag{11.2.14}$$

Taking the inverse of this last equation we get

$$V^{\dagger}\, V = 1\ . \tag{11.2.15}$$

So we have

$$V^{\dagger}\, V = V\, V^{\dagger} = 1\ . \tag{11.2.16}$$

Definition

An operator V is *unitary* if and only if it satisfies both equations (11.2.16). This definition is equivalent to the definition we gave in section 6.5 as we saw there. Combined with the calculation we have just done, it means that in order to transform from one basis to another, V must be a unitary operator. So, to change bases in $\mathcal{H}$ we require unitary transformations. Actually, the reason we must use unitary transformations is not as purely mathematical as the above argument would indicate. It is dictated by physics. The results of any physical measurement are contained in all the possible matrix elements or, what amounts to the same thing, inner products. This means that any "rotations" that we may perform must preserve all inner products. This is the same thing as requiring that the transformations be unitary.

11.2.1 Fourier Transform of Hermite Functions

As an example of what we just discussed we consider a unitary mapping from configuration space to momentum space where the basdis set in configurAation space is the hermite functions (simple harmonic oscillator eigenfunctions.

A complete orthonormal basis in $\mathcal{H}$ for one-dimensional problems is provided by the hermite functions $\{u_n(x)\}$ given by equation (9.2.47). If we choose units such that $k/\hbar\omega = 1$ then we have

$$u_n(x) = (-1)^n \frac{1}{\sqrt{n!}} \frac{1}{2^{n/2}} \left(\frac{d}{dx} - x \right)^n \frac{1}{\pi^{1/4}} e^{-x^2/2} . \tag{11.2.17}$$

These functions, as stated, form a complete orthonormal set. If we also admit continuous eigenvalues (rigged Hilbert space) then the functions

$$v_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx} \tag{11.2.18}$$

also form a complete orthonormal set. We now find the unitary operator V with matrix elements $V_{k,n}$ connecting the two sets. Notice that the index n is discrete while the index k is continuous. Thus,

$$v_k(x) = \sum_n V_{kn}\, u_n(x) \quad . \tag{11.2.19}$$

Using the orthonormality of the un we get

$$(v_k, u_m) = \sum_n V^*_{kn}\, (u_n, u_m) = V^*_{km} \, . \tag{11.2.20}$$

We therefore need only evaluate the inner product (v_k, u_m). We leave it as an exercise (problem 11.1) to show that the result is

$$V_{kn} = i^n\, u_n(k) \quad . \tag{11.2.21}$$

This shows furthermore that the Fourier transform of a hermite function is again a hermite function.

11.3 Dirac Notation

So far we have always, in a sense, used a representation of our states in only one definite basis, the position basis. This basis has, however, been suppressed in all cases. This characterization is really true only in a rigged Hilbert space but we continue to use the language of ordinary Hilbert space as explained at the end of Chapter 8. To make clear what is meant by the statement above, recall that the eigenfunctions of position are δ-functions. Thus,

$$x\delta(x-a) = a\delta(x-a) \ . \tag{11.3.22}$$

These form a complete orthonormal set and any function $\psi(x)$ can be considered as an expansion in terms of this set.

$$\psi(x) = \int_{-\infty}^{\infty} \psi(a)\,\delta(x-a)\,da \ . \tag{11.3.23}$$

This is not as purely formal as it seems and Dirac [11.2] very early devised an ingenious notation to take advantage of this. We now explain this notation. To begin with we consider an abstract vector space of states. Thus, linear superposition of states is defined but not an inner product. The elements of the space are called *kets* and are denoted by $|\cdot\rangle$. If we want to specify a specific ket we insert a label $|a\rangle$. This specification of kets is completely *basis independent.* This means that we do not explicitly write any wavefunction $\psi_a(x)$ but only write the symbol $|a\rangle$. Now just as in the case of $\psi_a(x)$, the label "a" usually refers to the eigenvalue of some operator A. This statement we now no longer write as

$$A\,\psi_a(x) = a\,\psi_a(x) \tag{11.3.24}$$

but as

$$A\,|a\rangle = a\,|a\rangle \ . \tag{11.3.25}$$

Both equations (11.3.24) and (11.3.25) say the same thing, except that in equation (11.3.25) we have not committed ourselves to a definite function or its Fourier transform or what have you. Corresponding to the space of kets we introduce the dual space (see section 8.1) of continuous linear functionals defined on the kets. This is called the space of *bras* and they are written $\langle\cdot|$. Specific bras are labelled in the same manner as kets. Furthermore if A is an operator on the space of kets then the corresponding operator on the space of bras is $A^\dagger$. Thus, the equation corresponding to (11.3.25) is given on the space of bras by

$$\langle a|A^\dagger = a^*\langle a| \ . \tag{11.3.26}$$

It corresponds to taking the dagger of equation (11.3.24), namely

$$\psi_a^*\,A^\dagger = a^*\,\psi_a^*(x) \ . \tag{11.3.27}$$

Since the bras are linear functionals over the kets they give a mapping from the kets into the complex numbers. We write this as $\langle\cdot|\cdot\rangle$ or for two specific ones as $\langle a|b\rangle$. (This makes it clear where the names "bra" and "ket" come from.) Again in terms of wavefunctions the corresponding expression is (ψ_a, ψ_b). The completeness relation for these states, as we verify later, reads

$$1 = \sum_n |n\rangle\langle n| \tag{11.3.28}$$

if the label n is discrete or

$$1 = \int k\rangle dk\langle k| \tag{11.3.29}$$

if the index k is continuous. These equations correspond to the equations

$$\delta(x-y) = \sum_n \psi_n(x)\psi_n^*(y) \tag{11.3.30}$$

if the index n is discrete or

$$\delta(x-y) = \int dk\psi_k(x)\psi_k^*(y) \tag{11.3.31}$$

if the index k is continuous.

The expectation value of an operator is now written as $\langle a|A|a\rangle$. It is important to remember that, although this expression looks symmetric with respect to left and right, the operator A *acts to the right* in this formula. The operator acting to the left is $A^\dagger$. Again in terms of wave functions this is written as $(\psi_a, A\psi_b)$ and clearly if A acts to the left we have $(A^\dagger\psi_a, \psi_b)$. In the compressed notation of Dirac we have to remember that $A^\dagger$ *acts to the left.* To establish precisely the connection between Dirac's bra, ket notation and the usual wavefunction formalism consider a specific eigenket $|n\rangle$ of the hamiltonian H. Thus,

$$H\,|n\rangle = E_n\,|n\rangle \ . \tag{11.3.32}$$

Now let the ket $|x\rangle$ be an eigenket of the position operator x_{op} so that

$$x_{op}\,|x\rangle = x\,|x\rangle \ . \tag{11.3.33}$$

We have written x_{op} but clearly this operator is just "multiplication by x". We now state that *the eigenfunctions* $\phi_n(x)$ *of the hamiltonian* H *in configuration space* are given by

$$\phi_n(x) = \langle x|n\rangle \tag{11.3.34}$$

$$\phi_n^*(x) = \langle n|x\rangle \ . \tag{11.3.35}$$

The orthogonality relation of two eigenfunctions ϕ_n and ϕ_m now follows from the orthogonality of the ket $|n\rangle$ and the bra $\langle m|$, namely

$$\langle m|n\rangle = \delta_{mn} \tag{11.3.36}$$

and the closure condition (completeness relation)

$$\int |x\rangle dx\langle x| = 1 \ . \tag{11.3.37}$$

This last equation is the same as equation (11.3.29). Thus, consider

$$\begin{aligned}(\phi_m, \phi_n) &= \int \phi_m^*(x)\phi_n(x)\,dx \\ &= \int \langle m|x\rangle dx\langle x|n\rangle \\ &= \langle m|n\rangle = \delta_{mn} \ .\end{aligned} \tag{11.3.38}$$

To make contact between equation (11.3.32) and the usual form of the time-independent Schrödinger equation we use (11.3.33) as well as the fact that if p is the momentum operator then

$$p|x\rangle = \frac{\hbar}{i}\frac{d}{dx}|x\rangle \ . \tag{11.3.39}$$

Then if

$$H = \frac{p^2}{2m} + V(x) \tag{11.3.40}$$

we take the inner product of equation (11.3.32) with the bra $\langle x|$ to get

$$\langle x|H|n\rangle = E_n\langle x|n\rangle = E_n\phi_n(x) \ . \tag{11.3.41}$$

The left side of this equation can be rewritten as follows:

$$\langle x|H|n\rangle = \int \langle x|H|y\rangle dy\langle y|n\rangle = \int \langle x|H|y\rangle dy\phi_n y) \ . \tag{11.3.42}$$

But using the explicit form for the Hamiltonian we have

$$\langle x|V(x)|y\rangle = V(y)\langle x|y\rangle = V(y)\delta(x-y) \tag{11.3.43}$$

where we have used the fact that $|y\rangle$ is an eigenket of the position operator x and hence also of $V(x)$. Furthermore,

$$\begin{aligned}\langle x|\frac{p^2}{2m}|y\rangle &= -\frac{\hbar^2}{2m}\langle x|\frac{d^2}{dy^2}|y\rangle \\ &= -\frac{\hbar^2}{2m}\frac{d^2}{dy^2}\langle x|y\rangle \\ &= -\frac{\hbar^2}{2m}\frac{d^2}{dy^2}\delta(x-y) \ .\end{aligned} \tag{11.3.44}$$

Substituting equations (11.3.43) and (11.3.44) into equation (11.3.41) and integrating by parts twice (or using the definition of the derivative of a distribution) we get:

$$\begin{aligned}\langle x|H|y\rangle &= \int \left[-\frac{\hbar^2}{2m}\frac{d^2\phi_n(y)}{dy^2} + V(y)\phi_n(y)\right]\delta(x-y)\,dy \\ &= -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\phi_n(x) + V(x)\phi_n(x)\end{aligned} \tag{11.3.45}$$

so that equation (11.3.32) now reads explicitly

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\phi_n(x)+V(x)\phi_n(x)=E_n\phi_n(x) \tag{11.3.46}$$

which is the usual form of the Schrödinger equation.

We have carried out these computations with excessive detail to illustrate all the steps that are involved. In practice, one frequently treats the bras and kets as if they were nothing more than wavefunctions with a little triangular bracket stuck on to them. This is all right in most circumstances; however, on occasion it is important to remember how they are really related to the wavefunction. This latter circumstance occurs when products of operators are operating on a ket. For example, if we have the expression $AB|\eta\rangle$ and we wish to write it in terms of wavefunctions then we can use the basis set $\{|x\rangle\}$ consisting of the eigenkets of the position operator. The expression is then rewritten as follows

$$\begin{aligned}\langle x|AB|\eta\rangle &= \int\langle x|A|y\rangle dy\langle y|B|z\rangle dz\langle z|\eta\rangle \\ &= \int dy\, dz\langle x|A|y\rangle\langle y|B|z\rangle\phi_\eta(z)\ .\end{aligned} \tag{11.3.47}$$

In a few paragraphs we shall see that this expression has a quite simple interpretation.

We next verify the completeness relations (11.3.28). The expression (11.3.29) is obtained in exactly the same way. Let $\{u_n\}$ be a complete basis and let $\{|n\rangle\}$ and $\{\langle n|\}$ denote the corresponding complete sets of kets and bras respectively so that

$$u_n(x)=\langle x|n\rangle \tag{11.3.48}$$

and

$$u_n^*(x)=\langle n|x\rangle\ . \tag{11.3.49}$$

Now any wavefunction ψ and corresponding ket $|\psi\rangle$ can be expanded as

$$\psi=\sum_n a_n u_n \tag{11.3.50}$$

or

$$|\psi\rangle=\sum_n a_n|n\rangle\ . \tag{11.3.51}$$

In both cases we have

$$a_n=(u_n,\psi)=\langle n|\psi\rangle\ . \tag{11.3.52}$$

Thus,

$$|\psi\rangle=\sum_n |n\rangle\langle n|\psi\rangle \tag{11.3.53}$$

which implies

$$\sum_n |n\rangle\langle n| = 1\,. \tag{11.3.54}$$

This is the completeness relation expressed in bra, ket notation.

If we now consider matrix elements of any operator A between wavefunctions ϕ and ψ we can write the whole expression in bra and ket notation. Using (11.3.53) and (11.3.54) we see that the matrix element $(\phi, A\psi)$ can be written as

$$\langle\phi|A|\psi\rangle = \sum_{mn}\langle\phi|m\rangle\langle m|A|n\rangle\langle n|\psi\rangle\,. \tag{11.3.55}$$

Then $\langle m|A|n\rangle$ is a "matrix" representing the operator A in the standard basis we have used for labelling our bras and kets. In fact the right side of (11.3.55) is symbolically a matrix product where $\langle n|\psi\rangle$ are column and $\langle\phi|m\rangle$ row matrix elements. Similarly $\langle m|A|n\rangle$ are the elements of a square matrix. This also shows that the right hand side of equation (11.3.47) may be interpreted as multiplying from left to right the square "matrices" $\langle x|A|y\rangle$ and $\langle y|B|z\rangle$ with the "column matrix" $\langle z|\eta\rangle = \phi_\eta(z)$.

Formulated in this way, quantum mechanics is historically referred to as matrix mechanics in contrast to the Schrödinger formulation which is called wave mechanics. Clearly they are just two different versions of the same thing, now called quantum mechanics. It is very useful to have many different formulations of the same theory. This allows us to choose the most convenient one for a particular computation and also gives us deeper insight into the structure of the theory. From now on we freely employ the Dirac notation and switch to different formulations as the mood or convenience strikes us. In fact, we shall soon lose track of exactly which formulation we are using because in Dirac notation no commitment to a particular formulation is required.

To begin the process of familiarizing ourselves with this notation we first revisit some aspects of the simple harmonic oscillator. After that we reconsider angular momentum and obtain matrix representations for the various operators.

11.4 Coherent States

Although the creation and annihilation operators, introduced when we studied the simple harmonic oscillator, are not self-adjoint it is still possible to diagonalize the annihilation operator. Clearly the ground state is an eigenstate of the annihilation operator corresponding to the eigenvalue zero. The eigenstates of the annihilation operator are known as *coherent states* and have found many applications in recent years in fields such as quantum optics. On the other hand, there are no normalizable eigenstates of the creation operator (see problem 11.13). This is because there is no state which will yield the ground state $|0\rangle$ after being acted on by the creation operator. On the other hand, all eigenstates of the number operator can be reached by action of the annihilation operator and so this operator can be diagonalized. The coherent states are the solutions of the equation

$$a|z\rangle = z|z\rangle \ . \tag{11.4.56}$$

To find these solutions we expand the state $|z\rangle$ in terms of the eigenstates $|n\rangle$ of the number operator $a^\dagger a$.

$$|z\rangle = \sum_n c_n(z)|n\rangle \ . \tag{11.4.57}$$

Applying the annihilation operator to this equation and using (11.4.56) we find

$$\sum_n c_n(z)\left[z|n\rangle - \sqrt{n}|n-1\rangle\right] = 0 \ . \tag{11.4.58}$$

Since the states $|n\rangle$ are linearly independent, this requires that

$$zc_{n-1} = \sqrt{n}c_n \tag{11.4.59}$$

or

$$c_n = \frac{z}{\sqrt{n}}c_{n-1} \ . \tag{11.4.60}$$

Therefore, we get after iteration that

$$c_n = \frac{z^n}{\sqrt{n!}}c_0 \ . \tag{11.4.61}$$

Substituting this into equation (11.4.57) and normalizing we find

$$|c_0|^2\, e^{|z|^2} = 1 \ . \tag{11.4.62}$$

Thus, we have the normalized eigenstates of the annihilation operator a.

$$|z\rangle = e^{-|z|^2/2}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}|n\rangle \ . \tag{11.4.63}$$

Notice that there are no restrictions on the eigenvalues z; they may be any complex number. Also the inner product between two different states is easily found to be (problem 11.11)

$$|\langle w|z\rangle|^2 = \exp - \left[|w|^2 + |z|^2 - w^* z - wz^*\right] = e^{-|w-z|^2} \ . \tag{11.4.64}$$

This shows that states belonging to different eigenvalues are not orthogonal. They do, however, form a complete set, in fact an overcomplete set. The corresponding formula is

$$1 = \frac{1}{\pi}\int d^2z\, |z\rangle\langle z| \ . \tag{11.4.65}$$

Here we are using that

$$z = x + iy \quad , \quad d^2z = dxdy \ . \tag{11.4.66}$$

The factor of $1/\pi$ in front of the integral shows that the set is overcomplete. To prove (11.4.65) we begin by expanding an arbitrary state $|f\rangle$ in terms of the eigenstates of the number operator

$$|f\rangle = \sum_n f_n |n\rangle \ . \tag{11.4.67}$$

This means that we have to show that

$$\begin{aligned} |f\rangle &= \sum_n f_n |n\rangle \\ &= \frac{1}{\pi}\int d^2z |z\rangle\langle z|f\rangle \\ &= \frac{1}{\pi}\sum_n f_n \int d^2z |z\rangle\langle z|n\rangle \ . \end{aligned} \tag{11.4.68}$$

We now insert (11.4.63) and change to polar coordinates using

$$z = r\,e^{i\theta} \quad , \quad d^2z = rdr\,d\theta \ . \tag{11.4.69}$$

Then we find that

$$\begin{aligned} \frac{1}{\pi}\sum_n f_n \int d^2z |z\rangle\langle z|n\rangle &= \frac{1}{\pi}\sum_{n,m} f_n \int rdr\,d\theta e^{-r^2}\, r^{n+m}\, e^{i(n-m)\theta}\frac{|m\rangle}{\sqrt{n!\,m!}} \\ &= \frac{1}{\pi}\sum_n f_n \int rdr\, e^{-r^2}\, r^{2n+1}\,\frac{|n\rangle}{n!} \\ &= \sum_n f_n |n\rangle = |f\rangle \ . \end{aligned} \tag{11.4.70}$$

This proves our result. Here we have used that

$$\frac{1}{2\pi}\int_0^{2\pi} d\theta\, e^{i(n-m)\theta} = \delta_{n,m} \tag{11.4.71}$$

and

$$\int_0^{\infty} dr\, r^{2n+1}\, e^{-r^2} = \frac{n!}{2} \ . \tag{11.4.72}$$

There is another convenient way of writing these coherent states. To do this we use the explicit form for the states $|n\rangle$.

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}\,|0\rangle \tag{11.4.73}$$

and substitute this in (11.4.63) to get

$$\begin{aligned} |z\rangle &= e^{-|z|^2/2}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle \\ &= e^{-|z|^2/2}\sum_{n=0}^{\infty}\frac{z^n(a^\dagger)^n}{n!}|0\rangle \\ &= e^{-|z|^2/2}\,e^{za^\dagger}|0\rangle \ . \end{aligned} \tag{11.4.74}$$

Finally we may also view any coherent state as the ground state of a new set of annihilation and creation operators, namely

$$b = a - z \quad \text{and} \quad b^\dagger = a^\dagger - z^* \tag{11.4.75}$$

since we clearly have

$$[b, b^\dagger] = 1 \tag{11.4.76}$$

and

$$b|z\rangle = 0 \ . \tag{11.4.77}$$

In configuration space the coherent states are just displaced ground states of the simple harmonic oscillator. To see this we use the explicit form of the annihilation operator

$$a = \sqrt{\frac{m\omega}{2\hbar}}x + \sqrt{\frac{\hbar}{2m\omega}}\frac{d}{dx} \tag{11.4.78}$$

and a wavefunction for a displaced ground state of a simple harmonic oscillator

$$\psi(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}\left(x - x_0\right)^2\right) \ . \tag{11.4.79}$$

Then, we find that

$$a\,\psi = \sqrt{\frac{m\omega}{2\hbar}}\,x_0\,\psi \ . \tag{11.4.80}$$

11.4.1 The Forced Simple Harmonic Oscillator

We now show that the ground state of a simple harmonic oscillator driven by a constant force F is just a coherent state. The Hamiltonian in this case is

$$\begin{aligned} H &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 - xF \\ &= \hbar\omega(a^\dagger a + 1/2) - \sqrt{\frac{\hbar}{2m\omega}}F(a^\dagger + a) \\ &= \hbar\omega\left(a^\dagger - \frac{F}{\sqrt{2m\hbar\omega^3}}\right)\left(a - \frac{F}{\sqrt{2m\hbar\omega^3}}\right) + \frac{1}{2}\hbar\omega - \frac{F^2}{2m\omega^2} \ . \end{aligned} \tag{11.4.81}$$

If we now define

$$\begin{aligned} b &= a - \frac{F}{\sqrt{2m\hbar\omega^3}} \\ b^\dagger &= a^\dagger - \frac{F}{\sqrt{2m\hbar\omega^3}} \end{aligned} \tag{11.4.82}$$

we see that this Hamiltonian is diagonalized with eigenvalues

$$E_n = (n + 1/2)\hbar\omega - \frac{F^2}{2m\omega^2} \ .$$

This also shows that the ground state is the coherent state $|F/\sqrt{2m\hbar\omega^3}\rangle$ since the equation

$$b\,|\frac{F}{\sqrt{2m\hbar\omega^3}}\rangle = 0 \tag{11.4.83}$$

implies that

$$a\,|\frac{F}{\sqrt{2m\hbar\omega^3}}\rangle = \frac{F}{\sqrt{2m\hbar\omega^3}}\,|\frac{F}{\sqrt{2m\hbar\omega^3}}\rangle \tag{11.4.84}$$

as claimed.

11.5 Quasi-classical States

We want to find a quantum state $|z\rangle$ that gives us the the best approximation for a classical simple harmonic oscillator. This means that we want a state in which the uncertainties Δx and Δp are constant in time and minimized. This problem is easiest to treat in the Heisenberg picture where, as we saw, the time-dependent annihilation and creation operators are given by

$$\begin{aligned} a(t) &= a\,e^{-i\omega t} \\ a^\dagger(t) &= a^\dagger\,e^{i\omega t}\,. \end{aligned} \tag{11.5.85}$$

Next we introduce the shifted operators

$$\begin{aligned} b(t) &= a(t) - \langle z|a(t)|z\rangle \\ b^\dagger(t) &= a^\dagger(t) - \langle z|a^\dagger(t)|z\rangle\,. \end{aligned} \tag{11.5.86}$$

Then,

$$\begin{aligned} (\Delta x(t))^2 &= \frac{\hbar}{2m\omega}\left[(\Delta b(0))^2\,e^{-2i\omega t} + (\Delta b^\dagger(0))^2\,e^{2i\omega t}\right. \\ &\quad \left.+\langle z|b^\dagger(0)b(0) + b(0)b^\dagger(0)|z\rangle\right] \\ (\Delta p(t))^2 &= -\frac{\hbar\omega m}{2}\left[(\Delta b(0))^2\,e^{-2i\omega t} + (\Delta b^\dagger(0))^2\,e^{2i\omega t}\right. \\ &\quad \left.-\langle z|b^\dagger(0)b(0) + b(0)b^\dagger(0)|z\rangle\right]\,. \end{aligned} \tag{11.5.87}$$

For these expressions to be time-independent requires that

$$\begin{aligned} (\Delta b(0))^2 &= \langle z|b(0)b(0)|z\rangle = 0 \\ (\Delta b^\dagger(0))^2 &= \langle z|b^\dagger(0)b^\dagger(0)|z\rangle = 0\,. \end{aligned} \tag{11.5.88}$$

Therefore, since there are no normalizable eigenstates of $b^\dagger(0)$ (see problem 11.13) we require that

$$\begin{aligned} b(0)|z\rangle &= 0 \\ \langle z|b^\dagger(0) &= 0\,. \end{aligned} \tag{11.5.89}$$

In that case we have that

$$\begin{aligned}(\Delta x(t))^2 &= \frac{\hbar}{2m\omega}\langle z|b(0)b^\dagger(0)|z\rangle \\ (\Delta p(t))^2 &= \frac{\hbar\omega m}{2}\langle z|b(0)b^\dagger(0)|z\rangle \end{aligned} \tag{11.5.90}$$

and

$$\begin{aligned}\Delta x(t)\Delta p(t) &= \frac{\hbar}{2}\langle z|b(0)b^\dagger(0)|z\rangle \\ &= \frac{\hbar}{2}\langle z|(a-z)(a^\dagger - z^*)|z\rangle \\ &= \frac{\hbar}{2}\left[\langle z|aa^\dagger|z\rangle + |z|^2\right] \\ &= \frac{\hbar}{2}\left(2|z|^2+1\right) \quad . \end{aligned} \tag{11.5.91}$$

This is clearly minimized if

$$|z| = \langle z|a|z\rangle = 0 \; . \tag{11.5.92}$$

This is therefore the best approximation for a classical state for the simple harmonic oscillator.

To see all of the above more clearly we now look at this result in configuration space where we are interested in the time-evolution of the displaced oscillator

$$\Psi(x,0) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-(m\omega/2\hbar)(x-x_0)^2} \; . \tag{11.5.93}$$

This is, as we saw earlier, a coherent state which may be written $\langle x|z\rangle$ where the parameter z labelling the coherent state is

$$z = \sqrt{\frac{m\omega}{2\hbar}}\, x_0 \quad .$$

This state can also be written as

$$\Psi(x,0) = \langle x|z\rangle = \langle x|e^{-ipx_0/\hbar}|0\rangle \quad . \tag{11.5.94}$$

Here $|0\rangle$ is the ground state of the simple harmonic oscillator. Also the state $\Psi(x,t)$ is the time evolved state

$$\Psi(x,t) = \langle x|e^{-iHt/\hbar}|z\rangle = \langle x|e^{-iHt/\hbar}\, e^{ipx_0/\hbar}|0\rangle. \tag{11.5.95}$$

Now,

$$e^{ipx_0/\hbar}|0\rangle = e^{-m\omega/4\hbar\, x_0^2}\, e^{\sqrt{m\omega/2\hbar}\; x_0 a^\dagger}|0\rangle \; . \tag{11.5.96}$$

Therefore,

$$\Psi(x,t) = e^{-m\omega/4\hbar\, x_0^2}\langle x|e^{-i(a^\dagger a+1/2)\omega t}\, e^{\sqrt{m\omega/2\hbar}\; x_0 a^\dagger}|0\rangle \; . \tag{11.5.97}$$

Expanding the exponential operators we find

$$\begin{aligned} & e^{-ia^\dagger a\omega t}\, e^{\sqrt{m\omega/2\hbar}\,x_0 a^\dagger}|0\rangle \\ = & \sum_{n,l=0}^{\infty} \frac{(-i\omega t)^n}{n!}(a^\dagger a)^n \frac{(\sqrt{m\omega/2\hbar}\,x_0)^l}{l!}(a^\dagger)^l|0\rangle\,. \end{aligned} \tag{11.5.98}$$

But, (see problem 11.14)

$$[a^\dagger a, (a^\dagger)^l] = l\,(a^\dagger)^l\,. \tag{11.5.99}$$

Therefore,

$$(a^\dagger a)^n (a^\dagger)^l|0\rangle = l^n (a^\dagger)^l|0\rangle \tag{11.5.100}$$

and

$$\begin{aligned} & e^{-ia^\dagger a\omega t}\, e^{\sqrt{m\omega/2\hbar}\,x_0 a^\dagger}|0\rangle \\ = & \sum_{n,l=0}^{\infty} \frac{(-i\omega t)^n}{n!}(l)^n \frac{(\sqrt{m\omega/2\hbar}\,x_0)^l}{l!}(a^\dagger)^l|0\rangle \\ = & \sum_{l=0}^{\infty} e^{-il\omega t}\frac{(\sqrt{m\omega/2\hbar}\,x_0)^l}{l!}(a^\dagger)^l|0\rangle \\ = & \sum_{l=0}^{\infty} \frac{(\sqrt{m\omega/2\hbar}\,x_0\,e^{-i\omega t})^l}{l!}(a^\dagger)^l|0\rangle \\ = & |\sqrt{m\omega/2\hbar}\,x_0\,e^{-i\omega t}\rangle \end{aligned} \tag{11.5.101}$$

where this last ket denotes a coherent state with the parameter

$$z = \sqrt{m\omega/2\hbar}\,x_0\,e^{-i\omega t}\quad . \tag{11.5.102}$$

So,

$$\begin{aligned} \Psi(x,t) = {} & \langle x|\sum_{n=0}^{\infty}\frac{(\sqrt{m\omega/2\hbar}\,x_0\,e^{-i\omega t})^n}{n!}(a^\dagger)^n|0\rangle\; e^{-m\omega/4\hbar\, x_0^2} \\ = {} & \sum_{n=0}^{\infty}\frac{(\sqrt{m\omega/2\hbar}\,x_0\,e^{-i\omega t})^n}{\sqrt{n!}}\langle x|n\rangle\; e^{-m\omega/4\hbar\, x_0^2} \\ = {} & \sum_{n=0}^{\infty}\frac{(\sqrt{m\omega/\hbar}\,x_0/2\,e^{-i\omega t})^n}{n!}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega}{2\hbar}x^2}\, H_n(\sqrt{m\omega/\hbar}\,x)\; e^{-m\omega/4\hbar\, x_0^2} \\ = {} & \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left\{-\left(\frac{m\omega}{2\hbar}\right)\left(x - x_0\,e^{-i\omega t}\right)^2\right\} \end{aligned} \tag{11.5.103}$$

where in the last step we have recognized the sum as the generating function for the hermite polynomials.

Now, if the parameter z in the expression for

$$|z\rangle = e^{-ipz}|0\rangle = e^{\sqrt{\frac{m\omega}{2\hbar}}za^\dagger}|0\rangle\, e^{-m\omega/4\hbar\, z^2} \tag{11.5.104}$$

were real then the state $|z\rangle$ would be properly normalized since e^{ipz} would be a unitary operator. However, if z is complex, e^{ipz} is no longer unitary and the state $|z\rangle$ has to be normalized. Thus, if for example $z = x_0 + iy_0$ then the proper normalization for

$$\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left\{-\left(\frac{m\omega}{2\hbar}\right)(x - x_0 - iy_0)^2\right\}$$

is

$$\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left\{-\left(\frac{m\omega}{2\hbar}\right)(x - x_0 - iy_0)^2\right\}\, e^{-m\omega/2\hbar\, y_0^2} .$$

Using this result we have that

$$\Psi(x,t) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\left(\frac{m\omega}{2\hbar}\right)(x - x_0\cos\omega t)^2}\, e^{i\left(\frac{m\omega}{2\hbar}\right)x_0 \sin\omega t(x - x_0\cos\omega t)} . \tag{11.5.105}$$

The probability density for this state is

$$\Psi(x,t) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/2} e^{-\left(\frac{m\omega}{\hbar}\right)(x - x_0\cos\omega t)^2} \tag{11.5.106}$$

and clearly shows that the peak of the displaced harmonic oscillator ground state wavefunction moves exactly like a classical simple harmonic oscillator.

11.6 Squeezed States

Equations (11.4.75) to (11.4.77) suggest a possible generalization of the coherent states by using another transformation that also preserves the commutation relations. The most general transformation has to be linear in a and $a^\dagger$. An example of this is

$$b = \frac{a - za^\dagger}{\sqrt{1 - |z|^2}} \quad , \quad b^\dagger = \frac{a^\dagger - z^* a}{\sqrt{1 - |z|^2}} \; . \tag{11.6.107}$$

This can still be followed by a transformation of the form given in (11.4.75) to yield the most general form. This transformation also preserves the commutation relations so that we still have

$$[b, b^\dagger] = 1 \; . \tag{11.6.108}$$

Equation (11.6.107) is an example of a *Bogoliubov* transformation. Also the state $|z\rangle$ now defined by

$$b|z\rangle = 0 \tag{11.6.109}$$

is again the ground state of the number operator $b^\dagger b$. To find the state $|z\rangle$ explicitly in terms of the eigenstates $|n\rangle$ of the number operator $a^\dagger a$ we proceed as before and write

$$|z\rangle = \sum_n c_{2n}(z)|2n\rangle \ . \tag{11.6.110}$$

In this case only even powers of n occur and so we have already made this explicit. Now writing out equation (11.6.109) using (11.6.107) and (11.6.110) we get

$$\sum_n c_{2n}(z)\left[\sqrt{2n}|2n-1\rangle - z\sqrt{2n+1}|2n+1\rangle\right] = 0 \ . \tag{11.6.111}$$

Therefore,

$$c_{2(n+1)} = z\sqrt{\frac{2n+1}{2n+2}}\,c_{2n} \ . \tag{11.6.112}$$

Iterating this equation we obtain

$$c_{2n} = z^n\sqrt{\frac{(2n-1)!!}{(2n)!!}}\,c_0 \tag{11.6.113}$$

where

$$(2n-1)!! = (2n-1)(2n-3)\ldots 5.3.1 \quad (-1)!! = 1 \tag{11.6.114}$$

$$(2n)!! = 2n.(2n-2)\ldots 6.4.2 = 2^n\,n! \quad 0!! = 1 \ . \tag{11.6.115}$$

The normalization of this state may be obtained by the following trick. We first compute the norm directly from equation (11.6.109) which may be written, using (11.6.107)

$$b|z\rangle \equiv (a - za^\dagger)|z\rangle = 0 \tag{11.6.116}$$

or

$$a|z\rangle = za^\dagger|z\rangle \ . \tag{11.6.117}$$

Taking the norm of this equation we find

$$\langle z|a^\dagger a|z\rangle = |z|^2\langle z|aa^\dagger|z\rangle = |z|^2\langle z|a^\dagger a + 1|z\rangle \ . \tag{11.6.118}$$

So that

$$\langle z|z\rangle = \frac{1-|z|^2}{z^2}\langle z|a^\dagger a|z\rangle \ . \tag{11.6.119}$$

Expanding both sides of this equation using (11.6.110) and (11.6.113) we get

$$|c_0|^2\sum_n |z|^{2n}\frac{(2n-1)!!}{(2n)!!} = \frac{1-|z|^2}{|z|^2}|c_0|^2\sum_n |z|^{2n}\frac{(2n-1)!!}{(2n-2)!!} \ . \tag{11.6.120}$$

If we call

$$\langle z|z\rangle = |c_0|^2 F(|z|) = |c_0|^2 \sum_n |z|^{2n} \frac{(2n-1)!!}{(2n)!!} \tag{11.6.121}$$

we see that (11.6.120) may be written

$$F = \frac{1-|z|^2}{|z|^2} |z| \frac{dF}{d|z|} . \tag{11.6.122}$$

Imposing the initial condition $F(0) = 1$ we find

$$F(|z|) = (1-|z|^2)^{-1/2} . \tag{11.6.123}$$

Therefore, we have

$$|z\rangle = (1-|z|^2)^{1/4} \sum_n z^n \sqrt{\frac{(2n-1)!!}{(2n)!!}} |2n\rangle . \tag{11.6.124}$$

This may be rewritten (see problem 11.11) as

$$|z\rangle = (1-|z|^2)^{1/4} \exp\left(\frac{z}{2} a^{\dagger 2}\right) |0\rangle \tag{11.6.125}$$

or in the even more suggestive form

$$|z\rangle = \exp\left(\frac{u}{2} a^{\dagger 2} - \frac{u^*}{2} a^2\right) |0\rangle \tag{11.6.126}$$

where

$$\tanh |u| = |z| . \tag{11.6.127}$$

The operator

$$S(u) = \exp\left(\frac{u}{2} a^{\dagger 2} - \frac{u^*}{2} a^2\right) \tag{11.6.128}$$

is called the *squeezing operator*. The reason for that is if u is real so that

$$u = |u| = r \tag{11.6.129}$$

or equivalently z is real

$$z = \tanh r \tag{11.6.130}$$

then, (see problem 11.11)

$$S^\dagger(r) x S(r) = \sqrt{\frac{\hbar}{2m\omega}} S^\dagger(r) a^\dagger S(r) = e^r x . \tag{11.6.131}$$

This clearly shows that the position variable is "squeezed" by the factor e^r so that

$$\langle z|x^2|z\rangle = e^{2r} \langle 0|x^2|0\rangle \tag{11.6.132}$$

where $|0\rangle$ is the ground state of the number operator $a^\dagger a$. Correspondingly it is easy to check that (see problem 11.12) that

$$\langle z|p^2|z\rangle = e^{-2r}\langle 0|p^2|0\rangle \ . \tag{11.6.133}$$

This shows that the coherent state $|z\rangle$ is still a minimum uncertainty state. Notice, however that in a squeezed state we can make the uncertainty in either x or p as small as we wish. This is what makes squeezed states so useful. This completes our discussion of coherent and squeezed states. A rather readable introduction to this subject is the article given in reference [11.5].

11.7 Example: Angular Momentum

Our whole discussion of angular momenta (if we do not restrict ourselves to orbital angular momenta) could have been carried out using only the algebraic properties of the operators. In this case it is conventional to denote the total angular momentum by J and the eigenvalues by j and m. Thus, the equations corresponding to (9.4.101) - (9.4.103) are

$$\begin{aligned} [J_x, J_y] &= i\hbar J_z & (11.7.134)\\ [J_y, J_z] &= i\hbar J_x & (11.7.135)\\ [J_z, J_x] &= i\hbar J_y \ . & (11.7.136)\end{aligned}$$

Also,

$$[J^2, \mathbf{J}] = 0 \ . \tag{11.7.137}$$

We again define

$$J_\pm = J_x \pm iJ_y \ . \tag{11.7.138}$$

The simultaneous eigenkets of J^2 , J_z are now denoted by $|j,m\rangle$. Thus, we have

$$J^2|j,m\rangle = j(j+1)\hbar^2|j,m\rangle \tag{11.7.139}$$

$$J_z|j,m\rangle = m\hbar|j,m\rangle \tag{11.7.140}$$

and corresponding to (9.5.176) we have

$$J_\pm|j,m\rangle = \sqrt{j(j+1)-m(m\pm 1)}\hbar|j,m\pm\rangle \ . \tag{11.7.141}$$

It is now a straightforward matter to evaluate the matrix elements of J^2, J_x, J_y and J_z. Thus,

$$\langle j',m'|J^2|j,m\rangle = j(j+1)\hbar^2\langle j',m'|j,m\rangle = j(j+1)\hbar^2\delta_{jj'}\delta_{mm'} \ . \tag{11.7.142}$$

So in this basis J^2 is given by a diagonal matrix. If the values of j, j' are fixed so that $j = j'$ then the matrix is $2j+1$ by $2j+1$. Thus for $j = j' = 1$

$$\langle 1,m'|J^2|1,m\rangle = 2\hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ . \tag{11.7.143}$$

Similarly, J_z is given by a diagonal matrix.

$$\langle 1,m'|J_z|1,m\rangle = m\hbar\delta_{jj'}\delta_{mm'} \ . \tag{11.7.144}$$

and for $j = j' = 1$ this becomes (recall problem 9.2)

$$\langle 1,m'|J_z|1,m\rangle = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \ . \tag{11.7.145}$$

To evaluate the matrix elements for J_x and J_y it is more convenient to first evaluate $J_\pm$. Thus,

$$\langle j',m'|J_\pm|j,m\rangle = \sqrt{j(j+1)-m(m\pm 1)}\,\hbar\delta_{jj'}\delta_{m,m'\pm 1} \ . \tag{11.7.146}$$

Again for $j = j' = 1$ these take the following form.

$$\langle 1,m'|J_+|1,m\rangle = \hbar \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \tag{11.7.147}$$

and

$$\langle 1,m'|J_-|1,m\rangle = \hbar \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \ . \tag{11.7.148}$$

They are obviously the hermitian conjugates of each other as is clear from

$$J_\pm = J_x \pm iJ_y \ . \tag{11.7.149}$$

We can now solve for

$$J_x = \frac{1}{2}(J_+ + J_-) \tag{11.7.150}$$

and

$$J_y = \frac{-i}{2}(J_+ - J_-) \tag{11.7.151}$$

to get

$$\langle 1,m'|J_x|1,m\rangle = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{11.7.152}$$

and

$$\langle 1,m'|J_y|1,m\rangle = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \ . \tag{11.7.153}$$

Both of these correspond to observables and so, as we see, are represented by hermitian matrices. The general scheme is given by (11.7.142), (11.7.143) and (11.7.146) and looks as follows

$$\langle j', m'|J^2|j, m\rangle$$

$$= \hbar^2 \left(\begin{array}{c|ccc|ccccc|c}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \text{etc.}
\end{array}\right) \tag{11.7.154}$$

for j, j' integral. The blocks are labelled by j j' and the elements within the blocks by m, m'.

We also have

$$\langle j', m'|J_z|j, m\rangle$$

$$= \hbar \left(\begin{array}{c|ccc|ccccc|c}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \text{etc.}
\end{array}\right) \tag{11.7.155}$$

and

$$\langle j', m' | J_- | j, m \rangle$$

$$= \hbar \left(\begin{array}{c|ccc|ccccc|c}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \sqrt{4} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \sqrt{6} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \sqrt{6} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{4} & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \text{etc.}
\end{array} \right) . \tag{11.7.156}$$

A specific example of the above general case (namely $j = 1/2$) has already been examined in detail in section 9.6. A brief re-reading of that section might be useful now. The matrices for $j = 1$ given by equations (11.7.144), (11.7.152) and (11.7.153) are completely equivalent to using the differential operators for L_x, L_y, L_z on the three-dimensional basis set $Y_{1,1}$, $Y_{1,0}$, $Y_{1,-1}$. In many instances it is much more convenient to use these matrix representations of the angular momentum operators rather than the differential operators.

11.8 Schrödinger Picture

The word "representation" is sometimes used in the literature with a somewhat different meaning from the one we have employed so far. We shall avoid that use and call these other representations "pictures" throughout this text. They occur when time-dependent unitary transformations are used. So far we have always worked in the so-called Schrödinger picture. In this picture the operators are generally time-independent and all the time dependence is carried by the states as given by the Schrödinger equation.

The matrix representations of operators in this picture, as stated, are time-independent unless they are *explicitly* time-dependent as in the case of an interaction that is switched on and off. We leave the discussion of such explicitly time-dependent operators for later (chapter 15).

Consider a basis set $\{u_n\}$ or $\{|n\rangle\}$. Then,

$$|\Psi(t)\rangle = \sum_n a_n(t)|n\rangle \tag{11.8.157}$$

and

$$i\hbar \frac{\partial |\Psi(t)\rangle}{\partial t} = H\,|\Psi(t)\rangle \tag{11.8.158}$$

can be rewritten to read

$$i\hbar\sum_n \frac{da_n(t)}{dt}|n\rangle = \sum_n a_n(t)H|n\rangle \ . \tag{11.8.159}$$

Or taking matrix elements with $\langle m|$ and using the fact that

$$\langle m|n\rangle = \delta_{m,n} \tag{11.8.160}$$

we get

$$i\hbar\frac{da_m(t)}{dt} = \sum_n H_{m,n}a_n(t) \tag{11.8.161}$$

where

$$H_{m,n} = \langle m|H|n\rangle \tag{11.8.162}$$

is clearly time independent if H does not explicitly depend on time. Similarly any other operator A has time independent matrix elements $\langle m|A|n\rangle$. Of course the expectation value of A is time dependent through the time dependence of $|\Psi(t)\rangle$.

11.9 Heisenberg Picture

We have previously seen (7.8.64) and (7.8.68) that the evolution of a state in the Schrödinger picture may be described by a unitary operator, the so-called evolution operator

$$U(t,t_0) = \exp\left[-\frac{iH}{\hbar}(t-t_0)\right] \ . \tag{11.9.163}$$

Suppose we apply the inverse of this unitary operator (with $t_0 = 0$) to every state of our Hilbert space. For concreteness let $|\Psi_S(t)\rangle$ be a state in the Schrödinger picture which has evolved according to $U(t) \equiv U(t,0)$ from $t = 0$. Then, we define a state in the Heisenberg picture $|\Psi_H\rangle$ by

$$\begin{aligned} |\Psi_H\rangle &= \exp\left(\frac{iH}{\hbar}t\right)|\Psi_S(t)\rangle \\ &= U^\dagger(t)|\Psi_S(t)\rangle \\ &= U^\dagger(t)U(t)|\Psi_S(0)\rangle \ . \end{aligned} \tag{11.9.164}$$

Thus,

$$|\Psi_H\rangle = |\Psi_S(0)\rangle \tag{11.9.165}$$

since $U^\dagger U = UU^\dagger = 1$ by virtue of the unitarity of the operator U. This shows that states $|\Psi_H\rangle$ in the Heisenberg picture are time-independent and coincide with states in the Schrödinger picture at time $t = 0$.

Since the physically interesting objects are expectation values or matrix elements of the type $\langle\Psi_S(t)|A|\Phi_S(t)\rangle$ these must remain invariant under our unitary transformation. Thus, it is necessary that the operators in the Schrödinger picture A_S transform to operators A_H in the Heisenberg picture. This gives

$$\langle\Psi_S(t)|A_S|\Phi_S(t)\rangle = \langle\Psi_H(t)|A_H|\Phi_H(t)\rangle \; . \tag{11.9.166}$$

Thus,

$$\begin{aligned}\langle\Psi_S(0)|U^\dagger(t)A_SU(t)|\Phi_S(t)\rangle &= \langle\Psi_H|U^\dagger(t)A_SU(t)|\Phi_H\rangle \\ &= \langle\Psi_H|A_H|\Phi_H\rangle \, .\end{aligned} \tag{11.9.167}$$

Hence,

$$A_H(t) = U^\dagger(t)A_SU(t) \, . \tag{11.9.168}$$

This shows that in the Heisenberg picture the time dependence is carried by the operators. To find the equation of motion they satisfy we simply use the fact that

$$i\hbar\frac{dU(t)}{dt} = HU(t) \tag{11.9.169}$$

and (remembering that $dH/dt = 0$) we differentiate (11.9.168) to get

$$\begin{aligned}i\hbar\frac{dA_H}{dt} &= i\hbar\frac{dU^\dagger}{dt}A_SU(t) + i\hbar U^\dagger A_S\frac{dU(t)}{dt} \\ &= -HU^\dagger A_SU + U^\dagger A_SUH \\ &= -HA_H + A_HH\end{aligned} \tag{11.9.170}$$

or

$$i\hbar\frac{dA_H}{dt} = [A_H, H] \, . \tag{11.9.171}$$

These are the famous *Heisenberg equations of motion.* They are identical with the classical Hamilton's equations if we replace $-i\hbar$ times the commutator (that is $-i\hbar[A, H]$) by the classical Poisson bracket $\{A, H\}$. We do not pursue this formal similarity any further, although this is the prescription originally used by Dirac to obtain quantization.

To illustrate this Heisenberg approach we consider once more the simple harmonic oscillator. Then,

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 \, . \tag{11.9.172}$$

In the Heisenberg picture the equations of motion are

$$i\hbar\frac{dx}{dt} = [x, H] = [x, \frac{p^2}{2m}] = i\hbar\frac{p}{m} \tag{11.9.173}$$

or

$$\dot{x} = \frac{p}{m} \tag{11.9.174}$$

and

$$i\hbar\frac{dp}{dt} = [p, H] = \frac{1}{2}m\omega^2[p, x^2] = -i\hbar m\omega^2 x \tag{11.9.175}$$

or

$$\dot{p} = -m\omega^2 x \quad . \tag{11.9.176}$$

These are clearly identical with the classical equations of motion. Hence,

$$\frac{d^2x}{dt^2} = \frac{\dot{p}}{m} = -\omega^2 x \tag{11.9.177}$$

$$\frac{d^2p}{dt^2} = -m\omega^2\dot{x} = -\omega^2 p\,. \tag{11.9.178}$$

Thus,

$$x = x_0 \cos\omega t + \frac{\dot{x}_0}{\omega}\sin\omega t \tag{11.9.179}$$

$$p = p_0 \cos\omega t + \frac{\dot{p}_0}{\omega}\sin\omega t \quad . \tag{11.9.180}$$

These solutions also appear identical with the classical ones; however, x_0, p_0, $\dot{x}_0$, $\dot{p}_0$ are now operators.

We furthermore have the first order equations

$$p = m\dot{x} = -m\omega x_0 \sin\omega t + m\dot{x}_0 \cos\omega t\,. \tag{11.9.181}$$

Therefore, using this condition at $t = 0$ we find

$$\frac{\dot{p}_0}{\omega} = -m\omega x_0 \tag{11.9.182}$$

and

$$p_0 = m\dot{x}_0 \tag{11.9.183}$$

so that

$$x = x_0 \cos\omega t + \frac{p_0}{m\omega}\sin\omega t \tag{11.9.184}$$

$$p = p_0 \cos\omega t - m\omega x_0 \sin\omega t \quad . \tag{11.9.185}$$

As already stated above, in spite of the similarity with the classical solutions, x_0 and p_0 are operators. In fact they are the same as the Schrödinger picture operators.

It is interesting to note that in this case also the creation and annihilation operators are useful. We have

$$\dot{x} = \frac{p}{m} \tag{11.9.186}$$

$$\dot{p} = -m\omega^2 x \; . \tag{11.9.187}$$

Introducing the annihilation and creation operators as defined by (9.2.4), (9.2.5) and (9.2.8), (9.2.9) we get

$$i\hbar\dot{a} = [a, H] = \hbar\omega a \tag{11.9.188}$$

$$i\hbar\dot{a}^\dagger = [a^\dagger, H] = -\hbar\omega a^\dagger \; . \tag{11.9.189}$$

The solutions are

$$a = a_0 e^{-i\omega t} \tag{11.9.190}$$

$$a^\dagger = a_0^\dagger e^{i\omega t} \; . \tag{11.9.191}$$

So in this case we had only first order equations to solve. The commutators in (11.9.188) and (11.9.189) were obtained from (9.2.16), (9.2.17). We could of course just have used (11.9.186) and (11.9.179) directly to get the same result

$$\begin{aligned} i\hbar\dot{a} &= i\hbar\frac{1}{\sqrt{2m\omega\hbar}}\left[m\omega\dot{x} + i\dot{p}\right] \\ &= i\hbar\frac{1}{\sqrt{2m\omega\hbar}}\left[\omega p - im\omega^2 x\right] \end{aligned} \tag{11.9.192}$$

or

$$i\hbar\dot{a} = \frac{\hbar\omega}{\sqrt{2m\omega\hbar}}\left[m\omega x + ip\right] = \hbar\omega a \tag{11.9.193}$$

which coincides with (11.9.188). This shows that introducing a, $a^\dagger$ would also simplify the classical problem for the simple harmonic oscillator. The states in this case are of course time independent and coincide with the states of the Schrödinger picture at time zero.

There is one more picture which is of great utility in applications and we discuss it next. It is clear, however, that it is possible to define as many different pictures as there are time-dependent unitary transformations. They vary, of course, in their utility. The so-called Interaction or Dirac picture is one of the more useful and has played a very important role in the development of Quantum Electrodynamics.

11.10 Dirac or Interaction Picture

As the name implies, this picture is useful in the case of interactions and can be thought of as lying half-way between the Schrödinger and Heisenberg pictures.

Suppose we have an, of necessity self-adjoint, Hamiltonian H which is itself the sum of two self adjoint operators H_0 and H' such that they have a common dense domain. That is,

$$H = H_0 + H' \; . \tag{11.10.194}$$

In practice H_0 will usually be an exactly diagonalizable Hamiltonian and H' a complicated interaction part. Again let $|\Psi_S(t)\rangle$ be a state in the Schrödinger picture and define the unitary operator

$$U_0(t) = \exp\left(-i\frac{H_0}{\hbar}t\right) \ . \tag{11.10.195}$$

The evolution of $|\Psi_S(t)\rangle$ is according to the Schrödinger equation with the full Hamiltonian

$$i\hbar\frac{\partial}{\partial t}|\Psi_S(t)\rangle = (H_0 + H')|\Psi_S(t)\rangle \ . \tag{11.10.196}$$

We then define a state in the Dirac picture $|\Psi_D(t)\rangle$ by

$$|\Psi_D(t)\rangle = U_0^\dagger(t)|\Psi_S(t)\rangle \ . \tag{11.10.197}$$

Since $U_0^\dagger(t)$ satisfies the equation

$$i\hbar\frac{\partial}{\partial t}U_0^\dagger = -H_0U_0^\dagger \ . \tag{11.10.198}$$

The equation of motion for $|\Psi_D(t)\rangle$ is found from

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\Psi_D(t)\rangle &= i\hbar\left(\frac{\partial}{\partial t}U_0^\dagger\right)|\Psi_S(t)\rangle + U_0^\dagger\frac{\partial}{\partial t}|\Psi_S(t)\rangle \\ &= -H_0U_0^\dagger|\Psi_S(t)\rangle + U_0^\dagger(H_0+H')|\Psi_S(t)\rangle \\ &= U_0^\dagger H'|\Psi_S(t)\rangle \\ &= U_0^\dagger H'U_0U_0^\dagger|\Psi_S(t)\rangle \end{aligned} \tag{11.10.199}$$

or

$$i\hbar\frac{\partial}{\partial t}|\Psi_D(t)\rangle = H'_D|\Psi_D(t)\rangle \tag{11.10.200}$$

where we have used $[U_0^\dagger, H_0] = 0$ and defined

$$H'_D = U_0^\dagger H'U_0 \ . \tag{11.10.201}$$

Thus, in the interaction picture the state evolves only according to the interaction part H'_D of the Hamiltonian. The price we have paid for this is that all operators *including* H'_D are now time dependent. In fact they satisfy an Heisenberg type of equation of motion. So, if A is any operator in the interaction picture, that is

$$A_D = U_0^\dagger A_SU_0 \ . \tag{11.10.202}$$

Then,

$$\begin{aligned} i\hbar\frac{d}{dt}A_D &= i\hbar\left(\frac{d}{dt}U_0^\dagger\right)A_SU_0 + i\hbar U_0^\dagger A_S\frac{d}{dt}U_0 \\ &= -H_0U_0^\dagger A_SU_0 + U_0^\dagger A_SU_0H_0 \end{aligned} \tag{11.10.203}$$

or

$$i\hbar\frac{d}{dt}A_D = [A_D, H_0] \ . \tag{11.10.204}$$

This is of course also the equation of motion for H'_D. Thus,

$$i\hbar\frac{d}{dt}H'_D = [H'_D, H_0] \ . \tag{11.10.205}$$

To see how all of this looks in matrix form consider solving first the equation

$$H_0|k\rangle = E_k|k\rangle \ . \tag{11.10.206}$$

Then, we can call

$$|k,t\rangle = |k\rangle e^{-iE_k t/\hbar} \tag{11.10.207}$$

and therefore write

$$|\Psi(t)\rangle = \sum_k a_k(t)\,|k\rangle e^{-iE_k t/\hbar} \tag{11.10.208}$$

where

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = H\,|\Psi(t)\rangle = (H_0 + H')|\Psi(t)\rangle \ . \tag{11.10.209}$$

Hence we get:

$$\begin{aligned} & i\hbar\sum_k \left(\dot{a}_k(t)\,|k\rangle e^{-iE_k t/\hbar} - i\frac{E_k}{\hbar}a_k(t)\,|k\rangle e^{-iE_k t/\hbar}\right) \\ = & \sum_k a_k(t)H_0|k\rangle e^{-iE_k t/\hbar} + \sum_k a_k(t)H'|k\rangle e^{-iE_k t/\hbar} \ . \end{aligned} \tag{11.10.210}$$

This simplifies using (11.10.206) to

$$i\hbar\sum_k \dot{a}_k(t)\,|k\rangle e^{-iE_k t/\hbar} = \sum_k a_k(t)H'|k\rangle e^{-iE_k t/\hbar} \ . \tag{11.10.211}$$

Multiplying by $\langle q|e^{iE_q t/\hbar}$ we get

$$H'_{q,k} = \langle q|H'|k\rangle \tag{11.10.212}$$

where

$$i\hbar\dot{a}_q = \sum_k H'_{q,k}e^{i[(E_q - E_k)/\hbar]t}a_k \tag{11.10.213}$$

and $e^{iE_q t/\hbar}H'e^{-iE_{k_q} t/\hbar}$ is the q,k matrix element of H' in the interaction picture and could be written as

$$H'_{D\,q,k} = e^{iE_q t/\hbar}H'e^{-iE_k t/\hbar} \ . \tag{11.10.214}$$

Thus,

$$i\hbar\frac{da_q}{dt} = \sum_k H'_{D\,q,k}a_k \ . \tag{11.10.215}$$

This is the matrix form of the Schrödinger equation in the interaction picture.

We return to these equations again when we discuss time-dependent perturbation theory.

11.11 Hidden Variables

We begin by recalling certain facts about the physical interpretation of quantum mechanics as discussed in chapter 7. If we have an observable $\mathcal{A}$, represented by a self-adjoint operator A , and a state $|\psi\rangle$ and we make a measurement of $\mathcal{A}$ then we must get one of the eigenvalues a_n of A as the result. Now, unless $|\psi\rangle$ is an eigenstate of A that is:

$$A|n\rangle = a_n|n\rangle \tag{11.11.216}$$

we cannot say which eigenvalue we will obtain. We can only give the probability

$$P_n = |\langle n|\psi\rangle|^2 \ . \tag{11.11.217}$$

In classical mechanics (for example statistical mechanics) situations also arise where we have only probabilities. This is due to the fact that, in these situations, we have an incomplete knowledge or specification of the state of the system of interest. In classical mechanics if the state of the system is completely specified then the value of any dynamical variable may be known with certainty. This has led people to speculate that the quantum mechanical specification of a state by a ket $|\psi\rangle$ may be incomplete and that there may be some, as yet unobserved, or hidden variables λ such that a complete specification say $|\psi\,,\lambda\rangle$ is possible. The λ's presumably are distributed with a classical probability $P(\lambda)$ so that $P(\lambda)d\lambda =$ fraction of particles with the variable λ lying between λ and $\lambda + d\lambda$. Also,

$$\int P(\lambda)d\lambda = 1 \quad , \quad P(\lambda) \geq 0 \ . \tag{11.11.218}$$

The result of a measurement (if the hidden variable λ is known) is now a definite number. The result of a measurement will be known with certainty. This, however, is not enough. We also have to recover the quantum mechanical results by averaging over the probability distribution of the hidden variables.

That such a program is doomed to failure was proven by J. S. Bell [11.3] in 1965. He obtained the remarkable result that any such hidden variable theory in which the hidden variables are local, i.e. do not affect each other over spacelike distances, must disagree with quantum mechanics. We now demonstrate his result with two spin 1/2 particles. The proof is, however, completely general since from any Hilbert space we can always extract such a finite dimensional subspace. We begin with some preliminaries. Consider two spin 1/2 particles, labelled 1 and 2, in a singlet ($s = 0$) state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[|+\rangle|-\rangle - |-\rangle|+\rangle\right] \ . \tag{11.11.219}$$

Here $|a\rangle|b\rangle$ represents the state in which particle 1 is in the state with spin a and particle 2 is in the state with spin b. Suppose these two particles now fly apart and we have two detectors set up far away to detect their spins. The detectors are set up so as to allow us to measure their spins along any one of three given directions specified by three unit vectors $\hat{n}_i$, $(i = 1, 2, 3)$. The corresponding spin operators are:

$$\mathbf{s} \cdot \hat{n}_i = \frac{\hbar}{2} \vec{\sigma} \cdot \hat{n}_i \quad . \tag{11.11.220}$$

For convenience, we factor out the $\hbar/2$ and consider only the quantities $\vec{\sigma} \cdot \hat{n}_i$. The angles between the three directions are

$$\begin{aligned} \hat{n}_1 \cdot \hat{n}_2 &= \cos\theta_{12} \\ \hat{n}_2 \cdot \hat{n}_3 &= \cos\theta_{23} \\ \hat{n}_3 \cdot \hat{n}_1 &= \cos\theta_{31} \quad . \end{aligned} \tag{11.11.221}$$

We can of course express the eigenstates $|\pm\rangle$ of spin up and down in terms of eigenstates of spin up and down along any of the three axes of the detectors namely the states $|i\ ,\ +\rangle$, $|i\ ,\ -\rangle$. Now suppose that we measure $\vec{\sigma} \cdot \hat{n}_i$ for particle 1 and $\vec{\sigma} \cdot \hat{n}_j$ for particle 2. We are interested in the probabilities of getting ± 1 for particle 1 and simultaneously ± 1 for particle 2. We call these probabilites:

$P_{i+,j+}$ = probability that particle 1 has spin $+\hbar/2$ along i and particle 2 has spin $+\hbar/2$ along j.

$P_{i+,j-}$ = probability that particle 1 has spin $+\hbar/2$ along i and particle 2 has spin $-\hbar/2$ along j.

$P_{i-,j+}$ = probability that particle 1 has spin $-\hbar/2$ along i and particle 2 has spin $+\hbar/2$ along j.

$P_{i-,j-}$ = probability that particle 1 has spin $-\hbar/2$ along i and particle 2 has spin $-\hbar/2$ along j.

We now compute these probabilities. The simplest way to do this is to align our coordinate system so that the vector $\hat{n}_i$ points along the z-axis.

$$\hat{n}_i = (0, 0, 1) \quad . \tag{11.11.222}$$

Then, $\vec{\sigma} \cdot \hat{n}_i = \sigma_3$ and the corresponding eigenstates are

$$|i, +\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{11.11.223}$$

$$|i, -\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad . \tag{11.11.224}$$

With $\hat{n}_i \cdot \hat{n}_j = \cos\theta_{ij}$ the most general possibility for $\hat{n}_j$ is

$$\hat{n}_j = (\sin\theta_{ij}\cos\alpha, \sin\theta_{ij}\sin\alpha, \cos\theta_{ij}) \quad . \tag{11.11.225}$$

In this case we have

$$\vec{\sigma} \cdot \hat{n}_j = \begin{pmatrix} \cos\theta_{ij} & e^{-i\alpha}\sin\theta_{ij} \\ e^{i\alpha}\sin\theta_{ij} & -\cos\theta_{ij} \end{pmatrix} \quad . \tag{11.11.226}$$

The eigenvalues of this matrix are, of course, ±1. The eigenstates are

$$|j\,,+\rangle = \begin{pmatrix} e^{-i\alpha/2}\cos(\theta_{ij}/2) \\ e^{i\alpha/2}\sin(\theta_{ij}/2) \end{pmatrix} \tag{11.11.227}$$

$$|j\,,-\rangle = \begin{pmatrix} e^{i\alpha/2}\sin(\theta_{ij}/2) \\ e^{-i\alpha/2}\cos(\theta_{ij}/2) \end{pmatrix} . \tag{11.11.228}$$

Now, recall that we are in a singlet state as in equation (11.11.219). This means

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[|i,+\rangle|j,-\rangle - |i,-\rangle|j,+\rangle\right] . \tag{11.11.229}$$

Therefore, the probability that electron 1 is in the state $|i\,,+\rangle$ is 1/2. Also if this is the case, then electron 2 is certainly in the state $|j\,,-\rangle$. This means that the probability $P_{i+,j+}$ is given by

$$P_{i+,j+} = \frac{1}{2}|\langle i,-|j,+\rangle|^2 \tag{11.11.230}$$

or

$$P_{i+,j+} = \frac{1}{2}\sin^2\frac{\theta_{ij}}{2} . \tag{11.11.231}$$

Similarly,

$$P_{i+,j-} = \frac{1}{2}|\langle i,-|j,\rangle|^2 = \frac{1}{2}\cos^2\frac{\theta_{ij}}{2} \tag{11.11.232}$$

$$P_{i-,j+} = \frac{1}{2}|\langle i,+|j,+\rangle|^2 = \frac{1}{2}\cos^2\frac{\theta_{ij}}{2} \tag{11.11.233}$$

$$P_{i-,j-} = \frac{1}{2}|\langle i,+|j,-\rangle|^2 = \frac{1}{2}\sin^2\frac{\theta_{ij}}{2} . \tag{11.11.234}$$

These are all the correlations we need. We can now show that these probabilities *cannot* be reproduced by hidden variables.

The idea is to perform the following experiments. We prepare the electrons in a state of total spin $s = 0$ and allow them to fly apart. Then far away, but equidistant from the source of electrons we have two spin detectors, A and B. For each detector we have three angle settings. The results for the two detectors are recorded and the correlations are then compared. During the experiment the angle setting is varied at random at each detector.

Now let us assume, contrary to quantum mechanics, that each electron already has definite spin properties when the two electrons separate and, not as in quantum mechanics, that their spin values emerge in the process of detection. This latter (quantum mechanical) assumption says that there is an instantaneous correlation between the two electrons no matter how far apart they are.

So, we assume that each electron carries six labels $(\lambda_1, \lambda_2, \lambda_3;\ \mu_1, \mu_2, \mu_3)$ where $\lambda_i = \pm 1$ and $\mu_j = \pm 1$. The λ_i are the values of $\vec{\sigma} \cdot \hat{n}_i$ measured by detector A on particle 1 and the μ_j are the values of $\vec{\sigma}\cdot\hat{n}_j$ measured by detector B on particle 2.

Let $P(\lambda_1,\lambda_2,\lambda_3;\mu_1,\mu_2,\mu_3)$ be the probabilities that for particle 1 in the directions $\hat{n}_i$ we get λ_i and at the same time for particle 2 in the directions $\hat{n}_j$ we get μ_j. These measurements are clearly correlated since the pair of electrons are in a state of total spin $s=0$. This means that

$$P(\pm,\lambda_2,\lambda_3;\pm,\mu_2,\mu_3)=0 \qquad (11.11.235)$$

$$P(\lambda_1,\pm,\lambda_3;\mu_1,\pm,\mu_3)=0 \qquad (11.11.236)$$

$$P(\lambda_1,\lambda_2,\pm;\mu_1,\mu_2,\pm)=0\ . \qquad (11.11.237)$$

In fact all the probabilities, except those with $\lambda_i=\mu_i$, vanish. Furthermore, all of these probabilities are non-negative. The space of hidden variables has (subject to constraints) $2^6=64$ parameters. Another constraint is

$$\sum P(\lambda_1,\lambda_2,\lambda_3;\mu_1,\mu_2,\mu_3)=1 \qquad (11.11.238)$$

where the sum runs over all values of the hidden variables.

We now use these expressions to recompute the probabilities

$$\begin{aligned} P_{1+,3+} &= \frac{1}{2}\sin^2\frac{\theta_{13}}{2} \\ &= \sum_{\lambda_2,\lambda_3;\mu_2,\mu_3} P(+,\lambda_2,\lambda_3;\mu_1,\mu_2,+) \\ &= \sum_{\lambda_2} P(+,\lambda_2,-;-,\lambda_2,+) \\ &= P(+,+,-;-,-,+)+P(+,-,-;-,+,+)\ . \end{aligned} \qquad (11.11.239)$$

Similarly,

$$\begin{aligned} P_{2+,3+} &= \frac{1}{2}\sin^2\frac{\theta_{23}}{2} \\ &= P(+,+,-;-,-,+)+P(-,+,-;+,-,+) \end{aligned} \qquad (11.11.240)$$

and

$$\begin{aligned} P_{1+,2+} &= \frac{1}{2}\sin^2\frac{\theta_{12}}{2} \\ &= P(+,-,+;-,+,-)+P(+,-,-;-,+,+)\ . \end{aligned} \qquad (11.11.241)$$

Now we add the last two equations and compare to the first to get:

$$\begin{aligned} & P(+,+,-;-,-,+)+P(-,+,-;+,-,+) \\ + & P(+,-,+;-,+,-)+P(+,-,-;-,+,+) \\ = & \frac{1}{2}\sin^2\frac{\theta_{23}}{2}+\frac{1}{2}\sin^2\frac{\theta_{12}}{2}+\frac{1}{2}\sin^2\frac{\theta_{13}}{2} \\ + & P(-,+,-;+,-,+)+P(+,-,+;-,+,-)\ . \end{aligned} \qquad (11.11.242)$$

Since all the probabilities are non-negative, this equation immediately yields one of Bell's inequalities

$$\frac{1}{2}\sin^2\frac{\theta_{23}}{2}+\frac{1}{2}\sin^2\frac{\theta_{12}}{2}\geq\frac{1}{2}\sin^2\frac{\theta_{13}}{2} \quad (11.11.243)$$

This inequality must hold if the hidden variables are to be able to reproduce the results of quantum mechanics. To check this, we now choose all three vectors $\hat{n}_i$ to lie in the same plane and such that $\hat{n}_2$ bisects the angle θ_{13}. So,

$$\theta_{12}=\theta_{23}=\frac{\theta_{13}}{2} \quad (11.11.244)$$

and we obtain

$$\sin^2\frac{\theta_{12}}{2}\geq\frac{1}{2}\sin^2\theta_{12}=2\sin^2\frac{\theta_{12}}{2}\cos^2\frac{\theta_{12}}{2} \ . \quad (11.11.245)$$

Hence we obtain

$$\cos^2\frac{\theta_{12}}{2}\leq\frac{1}{2} \quad (11.11.246)$$

or

$$\theta_{12}\geq\frac{\pi}{4} \ . \quad (11.11.247)$$

So, for the hidden variables theory to agree with quantum mechanics, this inequality must hold. If this inequality fails then the hidden variables can not reproduce the results of quantum mechanics. In this way we have a way to check experimentally the existence of hidden variables.

To try and bring out what is happening we now repeat this whole argument for a very specific case. For each particle there are three possible directions and two possible outcomes, namely ±1. Thus, for each particle there are $2^3 = 8$ possible parameters. Since $8\times 8 = 2^6$ we have the same result as previously, namely 64 parameters. The eight possible outcomes per particle are $+++$, $++-$, $+-+$, $-++$, $+--$, $-+-$, $--+$, $---$. These results are completely anticorrelated for the two particles since the total spin is zero. This means that if particle 1 is in the state $|+\rangle$ then particle 2 is in the state $|-\rangle$ and vice versa. We now consider two definite experiments.

a) The two detectors are exactly correlated. That is, we measure the outcomes for the settings 11, 22, 33. In this case we find perfect anticorrelation 100% of the time. The experiment yields only $+-$ or $-+$. Not a single case of $++$ or $--$ is observed.

b) We now choose the angles

$$\theta_{12}=\theta_{23}=\frac{\theta_{13}}{2}=\frac{\pi}{3} \ .$$

We chose this angle since it is greater than $\pi/4$ and thus will yield different results for quantum mechanics and hidden variables.

Now if the two detectors have different settings, that is 12, 13, 21, 23, 31, 32 we find that the experiment yields the following results. The detectors are anticorrelated exactly 1/4 of the time. That is, 1/8 of the time we measure $+-$ and 1/8 of the time we measure $-+$. The detectors are perfectly correlated 3/4 of the time. That is, 3/8 of the time we measure $++$ and 3/8 of the time we measure $--$. All of these results are in perfect agreement with quantum mechanics. To explain these results using a theory with hidden variables requires for part a) that the hidden variables for the two particles are completely anticorrelated. Thus, if particle 1 carries the information $++-$ then particle 2 must carry the information $--+$. In our previous language it means that the only non-zero probabilities occur for

$$P(\lambda_1, \lambda_2, \lambda_3; -\lambda_1, -\lambda_2, -\lambda_3) .$$

This assumption is required in order that we obtain agreement with the experimental results obtained in part a). We next turn our attention to the experiments in part b). Here we have the six possible detector settings already listed, namely 12, 21, 13, 31, 23, 32. Now suppose particles 1 and 2 carry the information $++-$ and $--+$ respectively. This is just the assumption necessary for part a) that the detectors are perfectly anticorrelated. In the case at hand, the detectors will now give opposite readings (anti-correlation) for the settings 12 and 21 but the same readings (correlation) for the settings 23, 32, 13, 31. So the hidden variable theory predicts that the readings will be anticorrelated exactly 1/3 of the time. The same result holds if any two labels are different, that is, the particles carry the information

$$\begin{array}{ll} +-+ & -+- \\ +-- & -++ \\ -++ & +-- \\ -+- & +-+ \\ --+ & ++- \end{array}$$

The only other possibilities are $+++$; $---$ and $---$; $+++$. In this latter case the theory always predicts perfect anticorrelation. Thus, with the hidden variable theory, we are led to the following conclusion. If the detectors are set at different angles, they will be anti-correlated at least 1/3 of the time. This is a specific case of Bell's inequality which reads

$$\text{The fraction of anticorrelations for different detector settings} \geq 1/3 \ .$$

As we saw, quantum mechanics predicts exactly 1/4 for the same result and is thus in violent disagreement with the hidden variable theory. Experiments performed by A. Aspect [11.4] agree with quantum mechanics at the level of three standard deviations. The accuracy of these experiments rules out any possibility of local hidden variable theories being able to replace quantum mechanics.

11.12 Problems

11.1 Verify equation (11.2.21), that is, compute the Fourier transform of the Hermite functions.

11.2 Find and solve the Heisenberg equation of motion for a particle with an Hamiltonian

$$H = \frac{p^2}{2m} \,.$$

If at $t = 0$ the particle is in the state $|0,0,0\rangle$ of a harmonic oscillator basis, find $\langle r^2 \rangle$ as a function of t.

11.3 Transform the displaced one-dimensional S.H.O.

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 + \beta x = H_0 + \beta x$$

to the Dirac picture. Solve for both $\psi_D(t)$, the wavefunction of the system and $x_D(t)$, the position operator of the system. Assume $\psi_D(0)$ describes the ground state of H_0 in the Schrödinger picture.
Hint: Rewrite everything in terms of the Schrödinger picture annihilation and creation operators and use the results of problem 9.14. The expression for $\psi_D(t)$ can not be completely evaluated.

11.4 Repeat problem 11.3 for the Heisenberg picture.

11.5 Use the Heisenberg picture to find the expectation values for p and x for a particle in an harmonic oscillator potential and also acted on by a constant force F. Assume that the state of the system is the ground state of the simple harmonic oscillator without the constant force F.

11.6 A free particle is acted on by a constant force F. Use the Heisenberg picture to find the expectation values of p and x if the state of the particle is described by a Gaussian wavefunction.

11.7 Repeat problem 11.6 in the Schrödinger picture.
Hint: Do not try to actually find the time-dependent wavefunction, but try to find an alternate method. It may help to review some of the results of Chapter 7. The point that this problem is illustrating is that in this case (a constant force) it is much simpler to do the calculation in the Heisenberg picture.

11.8 An electron in the potential $V(\mathbf{x})$ is also acted on by a weak constant magnetic field so that the resultant Hamiltonian is

$$H = \frac{1}{2m}\left(\mathbf{p} - e\mathbf{A}(\mathbf{x})\right)^2 + V(\mathbf{x})$$

with

$$\mathbf{A} = -\frac{1}{2}\mathbf{x} \times \mathbf{B} .$$

Introduce the interaction picture with

$$H_0 = \frac{1}{2m}(\mathbf{p})^2 + V(\mathbf{x})$$

and find the equations of motion for $\mathbf{p}$ and $\mathbf{x}$ in this picture. Write also the Schrödinger equation for the wavefunction in the interaction picture. Hint: Use the fact that the magnetic field is weak as well as the result that you will prove later (problem 17.8) that if $\mathbf{B}$ is a constant vector then

$$\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p} = \mathbf{B} \cdot \mathbf{L} .$$

11.9 Show that the normalized coherent state $|z\rangle$ in equation (11.4.74) may also be written

$$|z\rangle = e^{[za^\dagger - z^* a]}|0\rangle .$$

Hint: Use the results of problem 9.13.

11.10 Verify equation (11.4.64).

11.11 Verify equation (11.6.125).

11.12 Use the result of problem 9.15 to derive equation (11.6.131), namely

$$S^\dagger(r)xS(r) = \sqrt{\frac{\hbar}{2m\omega}}S^\dagger(r)(a + a^\dagger)S(r) = e^r x$$

as well as the result

$$S^\dagger(r)pS(r) = -i\sqrt{\frac{m\omega\hbar}{2}}S^\dagger(r)(a - a^\dagger)S(r) = e^{-r}p .$$

11.13 Show that the creation operator has no normalizable eigenstates. Hint: Assume that eigenstates exist and show that they are not normalizable.

11.14 Prove that

$$[a^\dagger a, (a^\dagger)^l] = l\,(a^\dagger)^l .$$

Hint: Use induction.

11.15 Given the Hamiltonian for a "forced" simple harmonic oscillator

$$H(t) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 - xF(t) - pG(t)$$

where both $F(t)$ and $G(t)$ vanish outside the interval $0 < t < T$, calculate the probability $P(n)$ that the system is in the state of n quanta of the Hamiltonian $H(t > T)$ if it was originally in the ground state of the Hamiltonian $H(t < 0)$.
Hint: Use the Heisenberg picture.

Bibliography

[11.1] The transformation theory is due in large measure to
J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton (1955).

[11.2] It is also of very great value to read:
P.A.M. Dirac, *Principles of Quantum Mechanics*, Oxford University Press, Oxford, 3rd edition (1947).

[11.3] J.S. Bell, Physics **1**,195 (1964).

[11.4] A. Aspect, Phys. Rev. **14D**, 1944 (1976).

[11.5] R. Munõz-Tapia, Am.J.Phys. **61**, 1005 (1993).

Chapter 12

Non-Degenerate Perturbation Theory

12.1 Introduction

So far we have only discussed problems that permitted exact analytic solutions. This is, in fact, seldom the case and one is forced to use approximation procedures. The exact solutions obtained so far then help to give one an intuitive feeling for what to expect in complicated situations. Thus, faced with what may be a more or less exact but complicated, Hamiltonian H describing a given physical situation one can adopt several different approaches.

In many situations it is convenient to view the Hamiltonian as a mathematical model of the situation. It then becomes reasonable to modify the model until one arrives at a simpler Hamiltonian that still retains the essential physical features. To check the model dependence of the results one can then construct many different models all of which, in some sense, approximate the situation of interest and study how the results vary from model to model. This approach requires a good deal of physical insight and has been much employed in nuclear and solid state physics.

On the other hand, even using model Hamiltonians one is frequently still left with a problem that cannot be solved exactly and such that further approximations on the model would destroy the physics of interest. In this case one requires techniques for handling such Hamiltonians in an approximate fashion. Here also the problems divide into essentially two classes. In one case one is interested in the modification of the stationary states of the system under the perturbation. In the second case one is interested not in the shifts in the stationary states but rather in the transition between stationary states due to the perturbation. This latter situation usually arises with time-dependent perturbations. We shall not consider the second case for several chapters but concentrate instead on the first case for now.

A particularly happy situation occurs if in a given physical problem, say an Hamiltonian H, it is possible to split the problem into a relatively simple (solvable) part H_0 and a "small" part $\lambda H'$ so that

$$H = H_0 + \lambda H' \ . \tag{12.1.1}$$

Small here means that in some sense H' and H_0 are comparable and $\lambda < 1$ or else $\lambda \approx 1$ and H' is small compared to H_0. Both these statements are vague since both H_0 and $\lambda H'$ are unbounded. To make them concrete, we assume for the time being that for any eigenket $|n\rangle$ of H_0

$$|\langle n|H_0|n\rangle| \geq |\langle n|\lambda H'|n\rangle| \, . \tag{12.1.2}$$

In such a case it is possible to treat $\lambda H'$ as a perturbation on H_0. A number of different and useful techniques have been developed for this and it is the purpose of this chapter to study some of these.

12.2 Rayleigh-Schrödinger Perturbation Theory

Consider an Hamiltonian of the form

$$H = H_0 + \lambda H' \tag{12.2.3}$$

where $\lambda H'$ is small with respect to H_0. We further suppose that H_0 is sufficiently simple that we are able to solve its eigenvalue problem exactly. Thus, we have

$$H_0|n\rangle^{(0)} = E_n^{(0)}|n\rangle^{(0)} \tag{12.2.4}$$

where $|n\rangle^{(0)}$, $E_n^{(0)}$ are the exact eigenkets and eigenvalues of H_0. We assume throughout this section that the eigenvalue of interest, say $E_n^{(0)}$, is non-degenerate. The reason for this assumption will become obvious shortly. We further assume that $E_n^{(0)}$ is a discrete eigenvalue and that as $\lambda \to 0$ the exact eigenvalue E_n of H approaches $E_n^{(0)}$. In this case it is reasonable that for small λ we can expand the eigenfunction $|n\rangle$ and eigenvalue E_n of H in a fairly rapidly converging power series in λ. The eigenket $|n\rangle$ is specified only up to an arbitrary normalization constant which we fix by requiring that

$$^{(0)}\langle n|n\rangle = \ ^{(0)}\langle n|n\rangle^{(0)} = 1 \ . \tag{12.2.5}$$

This assumption states that all "corrections" to the unperturbed state $|n\rangle^{(0)}$ are orthogonal to $|n\rangle^{(0)}$ and not merely a change in the normalization. Also our assumption of a power series now leads us to write

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots \tag{12.2.6}$$

$$|n\rangle = |n\rangle^{(0)} + \lambda|n\rangle^{(1)} + \lambda^2|n\rangle^{(2)} + \ldots \ . \tag{12.2.7}$$

Substituting this in the Schrödinger equation

$$H|n\rangle = E_n|n\rangle \tag{12.2.8}$$

and equating the coefficients of equal powers of λ we get:

$$(H_0 - E_n^{(0)})|n\rangle^{(0)} = 0 \tag{12.2.9}$$

$$(H_0 - E_n^{(0)})|n\rangle^{(1)} + (H' - E_n^{(1)})|n\rangle^{(0)} = 0 \tag{12.2.10}$$

$$(H_0 - E_n^{(0)})|n\rangle^{(2)} + (H' - E_n^{(1)})|n\rangle^{(1)} - E_n^{(2)}|n\rangle^{(0)} = 0 \tag{12.2.11}$$

$$\begin{aligned}&(H_0 - E_n^{(0)})|n\rangle^{(r)} + (H' - E_n^{(1)})|n\rangle^{(r-1)} - E_n^{(2)}|n\rangle^{(r-2)}\\ &- \ldots - E_n^{(r)}|n\rangle^{(0)} = 0\ .\end{aligned} \tag{12.2.12}$$

Clearly (12.2.9) is identical with (12.2.4) and simply states that $E_n^{(0)}$ and $|n\rangle^{(0)}$ are eigenvalues and eigenfunctions respectively of H_0. The normalization condition (12.2.5) when substituted into (12.2.7) now gives

$$^{(0)}\langle n|n\rangle^{(0)} = \ ^{(0)}\langle n|n\rangle^{(0)} + \lambda\ ^{(0)}\langle n|n\rangle^{(1)} + \lambda^2\ ^{(0)}\langle n|n\rangle^{(2)} + \ldots \tag{12.2.13}$$

so that we get

$$^{(0)}\langle n|n\rangle^{(1)} = \ ^{(0)}\langle n|n\rangle^{(2)} = \ldots = \ ^{(0)}\langle n|n\rangle^{(r)} = 0\ . \tag{12.2.14}$$

This simply reflects, as we stated above, the fact that we have chosen all corrections to $|n\rangle^{(0)}$ to be orthogonal to $|n\rangle^{(0)}$. To repeat, the corrections involve only the changes in $|n\rangle^{(0)}$ and not any rescaling of $|n\rangle^{(0)}$.

Now, (12.2.9) determines the zeroth order approximation for E_n and $|n\rangle$, (12.2.10) the first order approximation in terms of the zeroth and finally (12.2.12) the rth order approximation in terms of all the lower order approximations. To see this we simply form the inner product with $^{(0)}\langle n|$ of equation (12.2.12) to get

$$^{(0)}\langle n|H'|n\rangle^{(r-1)} - E_n^{(r)} = 0 \tag{12.2.15}$$

where we have used (12.2.14) and the fact that H_0 is self-adjoint. If we now further take the inner product of equation (12.2.12) with $^{(0)}\langle n|$ for $m \neq n$ we get the components of $|n\rangle^{(r)}$ along $|m\rangle^{(0)}$, and together with (12.2.14) this specifies $|n\rangle^{(r)}$ completely in terms of the lower order corrections. Thus, we get

$$\begin{aligned}&\left(E_m^{(0)} - E_n^{(0)}\right)\ ^{(0)}\langle m|n\rangle^{(r)} + ^{(0)}\langle m|H' - E_n^{1}|n\rangle^{(r-1)} - E_n^{(2)}\ ^{(0)}\langle m|n\rangle^{(r-2)}\\ &- \ldots - E_n^{(r-1)}\ ^{(0)}\langle m|n\rangle^{(1)} = 0 \qquad m \neq n\end{aligned} \tag{12.2.16}$$

or

$$\begin{aligned}&^{(0)}\langle m|n\rangle^{(r)} = \frac{1}{E_n^{(0)} - E_m^{(0)}}\Big[^{(0)}\langle m|H' - E_n^{1}|n\rangle^{(r-1)} - E_n^{(2)}\ ^{(0)}\langle m|n\rangle^{(r-2)}\\ &- E_n^{(3)}\ ^{(0)}\langle m|n\rangle^{(r-3)} - \ldots - E_n^{(r-1)}\ ^{(0)}\langle m|n\rangle^{(1)}\Big] \qquad m \neq n\ .\end{aligned} \tag{12.2.17}$$

We next look at the lowest orders of perturbation in more detail.

12.3 First Order Perturbations

If we set $r = 1$ in (12.2.15) and (12.2.17) we get that

$$E_n^{(1)} = {}^{(0)}\langle n|H'|n\rangle^{(1)} \tag{12.3.18}$$

and

$$\begin{aligned} {}^{(0)}\langle m|n\rangle^{(1)} &= \frac{1}{E_n^{(0)} - E_m^{(0)}}\, {}^{(0)}\langle m|H' - E_n^1|n\rangle^{(0)} \\ &= \frac{{}^{(0)}\langle m|H'|n\rangle^{(0)}}{E_n^{(0)} - E_m^{(0)}} \qquad m \neq n\ . \end{aligned} \tag{12.3.19}$$

Hence to first order in λ

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} = {}^{(0)}\langle n|H_0 + \lambda H'|n\rangle^{(0)}\ . \tag{12.3.20}$$

This is just the expectation value of the total Hamiltonian H in the unperturbed state $|n\rangle^{(0)}$.

Also from (12.3.19) and since ${}^{(0)}\langle n|n\rangle^{(1)} = 0$ we see that to first order in λ the eigenket

$$|n\rangle = |n\rangle^{(0)} + \lambda|n\rangle^{(1)} \tag{12.3.21}$$

is

$$|n\rangle = |n\rangle^{(0)} + \lambda \sum_{m\neq n} \frac{|m\rangle^{(0)}\, {}^{(0)}\langle m|H'|n\rangle^{(0)}}{E_n^{(0)} - E_m^{(0)}}\ . \tag{12.3.22}$$

We can also write this as

$$|n\rangle = |n\rangle^{(0)} + \lambda \sum_{m\neq n} \left(E_n^{(0)} - H_0\right)^{-1} |m\rangle^{(0)}\, {}^{(0)}\langle m|H'|n\rangle^{(0)}\ . \tag{12.3.23}$$

Since $|n\rangle^{(0)}$ is already normalized it might appear from (12.3.22) or (12.3.23) that $|n\rangle$ is not. However, to order λ, $|n\rangle$ is indeed normalized since

$$\begin{aligned} \langle n|n\rangle &= {}^{(0)}\langle n|n\rangle^{(0)} + \lambda^2 \sum_{m\neq n} \frac{{}^{(0)}\langle n|H'|m\rangle^{(0)}\, {}^{(0)}\langle m|H'|n\rangle^{(0)}}{\left(E_n^{(0)} - E_m^{(0)}\right)^2} \\ &= {}^{(0)}\langle n|n\rangle^{(0)} + \lambda^2 \left|\frac{H'_{n,m}}{E_n^{(0)} - E_m^{(0)}}\right|^2 \\ &= {}^{(0)}\langle n|n\rangle^{(0)} + \text{terms of order } \lambda^2\ . \end{aligned} \tag{12.3.24}$$

The cross-terms cancelled in view of the self-adjointness of H_0 and H'. Clearly the degree of convergence of $|n\rangle$ is determined by the ratio

$$\left|\frac{H'_{n,m}}{E_n^{(0)} - E_m^{(0)}}\right|^2\ . \tag{12.3.25}$$

It is clear now why the assumption of non-degeneracy was made. For otherwise, $E_n^{(0)} - E_m^{(0)}$ could vanish giving a divergent rather than a small finite result. In fact if H_0 has close-lying levels so that $E_n^{(0)} - E_m^{(0)}$ is small, these levels also have to be treated as if they were degenerate.

We next illustrate the above procedure by means of some examples.

12.4 Anharmonic Oscillator

The Hamiltonian is

$$\begin{aligned} H &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 + \lambda x^4 \\ &= H_0 + \lambda x^4 \ . \end{aligned} \tag{12.4.26}$$

We consider the first order perturbation of the ground state $|0\rangle$ of H_0. Thus,

$$H_0|0\rangle = \frac{1}{2}\hbar\omega|0\rangle \tag{12.4.27}$$

and in terms of the annihilation operator a

$$a|0\rangle = 0 \tag{12.4.28}$$

and

$$x = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger) \tag{12.4.29}$$

where we have used (9.2.5) and (9.2.12). Using (12.3.20) we get the ground state energy to first order in λ as

$$\langle 0|H_0 + \lambda x^4|0\rangle = \frac{1}{2}\hbar\omega + \lambda\langle 0|x^4|0\rangle \ . \tag{12.4.30}$$

But,

$$\begin{aligned} \langle 0|x^4|0\rangle &= \left(\frac{\hbar}{2m\omega}\right)^2 \langle 0|(a + a^\dagger)^4|0\rangle \\ &= \left(\frac{\hbar}{2m\omega}\right)^2 \langle 0|a^4 + a^3 a^\dagger + a^2 a^\dagger a + a^2 a^{\dagger 2} \\ &\quad + a a^\dagger a^2 + a a^\dagger a a^\dagger + a a^{\dagger 2} a + a a^{\dagger 3} + a^\dagger a^3 + a^\dagger a^2 a^\dagger \\ &\quad + a^\dagger a a^\dagger a + a^\dagger a a^{\dagger 2} + a^{\dagger 2} a^2 + a^{\dagger 2} a a^\dagger + a^{\dagger 3} a + a^{\dagger 4}|0\rangle \\ &= \left(\frac{\hbar}{2m\omega}\right)^2 \langle 0|a^2 a^{\dagger 2} + a a^\dagger a a^\dagger|0\rangle \\ &= 3\left(\frac{\hbar}{2m\omega}\right)^2 \ . \end{aligned} \tag{12.4.31}$$

Hence, to first order in λ the ground state energy is given by

$$E_0 = \frac{1}{2}\hbar\omega\left(1 + \frac{3\lambda}{2}\frac{\hbar}{m^2\omega^3}\right) \ . \tag{12.4.32}$$

12.5 Ground State of Helium-like Ions

The Hamiltonian in this case is

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{r_{12}} \tag{12.5.33}$$

where (see figure 12.1) $\mathbf{r}_1$ and $\mathbf{r}_2$ are the position vectors of the first and second electron respectively and $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ is the separation between the two electrons. Also $Z = 2$ for helium but we leave it arbitrary for the time being since we also consider helium-like ions. If we at first neglect the repulsion of the two electrons we get:

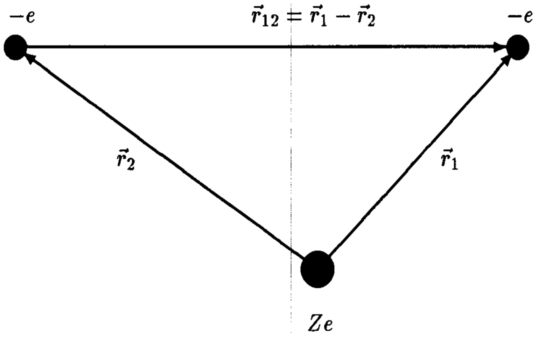

Figure 12.1: Geometry for helium-like ions.

$$H_0 = H_{01} + H_{02} \tag{12.5.34}$$

where

$$H_{0i} = \frac{p_i^2}{2m} - \frac{Ze^2}{r_i} \tag{12.5.35}$$

is just the hydrogenic Hamiltonian. Furthermore,

$$[H_{01}, H_{02}] = 0 \ . \tag{12.5.36}$$

Therefore we can diagonalize H_{01} and H_{02} and hence $H_0 = H_{01} + H_{02}$ simultaneously. The solution for the ground state energy is then

$$E_0^{(0)} = -2\frac{1}{2}\frac{Ze^2}{a} \tag{12.5.37}$$

where

$$a = \frac{a_0}{Z} = \frac{\hbar^2}{Zme^2} \tag{12.5.38}$$

and the ground state eigenfunction of H_0 is

$$\psi_0^{(0)}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{\pi a^3} e^{-(r_1+r_2)/a} \ . \tag{12.5.39}$$

The correction $E_0^{(1)}$ due to the electrostatic repulsion of the two electrons is then, to first order, given by

$$\begin{aligned} E_0^{(0)} &= \int \psi_0^{(0)*} \frac{e^2}{r_{12}} \psi_0^{(0)} \, d^3r_1 \, d^3r_2 \\ &= \frac{e^2}{\pi^2 a^6} \int \frac{e^{-2(r_1+r_2)/a}}{|\mathbf{r}_1 - \mathbf{r}_2|} d^3r_1 \, d^3r_2 \ . \end{aligned} \tag{12.5.40}$$

This is just the electrostatic energy due to two spherical charge distributions with charge densities

$$\rho(\mathbf{r}) = -\frac{e}{\pi a^3} e^{-2r/a} \ . \tag{12.5.41}$$

To evaluate the integrals in (12.5.40) we first expand $1/|\mathbf{r}_1 - \mathbf{r}_2|$ in terms of spherical harmonics (problem 12.1)

$$\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_{l,m}^*(\hat{\mathbf{r}}_1) Y_{l,m}(\hat{\mathbf{r}}_2) \tag{12.5.42}$$

where

$$r_< = \begin{cases} r_1 & \text{if} \quad r_1 < r_2 \\ r_2 & \text{if} \quad r_2 < r_1 \end{cases} \quad \text{and} \quad r_> = \begin{cases} r_2 & \text{if} \quad r_1 < r_2 \\ r_1 & \text{if} \quad r_2 < r_1 \end{cases} \ . \tag{12.5.43}$$

Then,

$$\begin{aligned} E_0^{(1)} &= 16\pi^2 \int_0^\infty \rho(r_1 r_1 dr_1 \Bigg(\int_0^{r_1} \rho(r_2 r_2^2 dr_2 + r_1 \int_{r_1}^\infty \rho(r_2 r_2 dr_2 \Bigg) \\ &= \frac{5}{4} Z \frac{1}{2} \frac{e^2}{a_0} \end{aligned} \tag{12.5.44}$$

where we have used the explicit form of $\rho(\mathbf{r})$ and (12.5.42). Hence the ground state energy to first order in the electrostatic repulsion of the two electrons is given by

$$E_0 = -\frac{e^2}{a_0} Z^2 \left(1 - \frac{5}{8Z} \right) \ . \tag{12.5.45}$$

Since the perturbation decreases with increasing Z we expect the accuracy of our result to improve as we go from He to Be $^{++}$. The results are as presented in table 12.1.

Table 12.1: Ground State Energies of Helium & Helium-Like Ions

Atom	Z	E_0 eV (calculated)	E_0 eV experimental	% Error
He	2	-74	-79	6
Li^+	3	-193	-197	2
Be^{++}	4	-366	-370	1

12.6 Second Order Perturbations

We now take a closer look at perturbations to second order. From (12.2.15) we get that

$$E_n^{(2)} = {}^{(0)}\langle n|H'|n\rangle^{(1)} . \tag{12.6.46}$$

But we also have that

$$|n\rangle^{(1)} = \sum_{m\neq n} \frac{|m\rangle^{(0)}\,{}^{(0)}\langle m|H'|n\rangle^{(0)}}{E_n^{(0)} - E_m^{(0)}} . \tag{12.6.47}$$

Hence we get

$$E_n^{(2)} = \sum_{m\neq n} \frac{{}^{(0)}\langle n|H'|m\rangle^{(0)}\,{}^{(0)}\langle m|H'|n\rangle^{(0)}}{E_n^{(0)} - E_m^{(0)}} . \tag{12.6.48}$$

Or calling, as before

$$H'_{n,m} = {}^{(0)}\langle n|H'|m\rangle^{(0)} \tag{12.6.49}$$

we have to second order in λ

$$E_n = E_n^{(0)} + \lambda\,{}^{(0)}\langle n|H'|m\rangle^{(0)} + \lambda^2 \sum_{m\neq n} \frac{|H'_{n,m}|^2}{E_n^{(0)} - E_m^{(0)}} . \tag{12.6.50}$$

We leave it as an exercise (problem 12.2) to show that by setting $r = 2$ in equation (12.2.17) one gets the second order correction to the wavefunction

$$\begin{aligned} |n\rangle^{(2)} &= \sum_{m,r\neq n} \frac{|m\rangle^{(0)}\,{}^{(0)}\langle m|H'|r\rangle^{(0)}\,{}^{(0)}\langle r|H'|n\rangle^{(0)}}{\left(E_n^{(0)} - E_m^{(0)}\right)\left(E_n^{(0)} - E_r^{(0)}\right)} \\ &- \sum_{m\neq n} \frac{|m\rangle^{(0)}\,{}^{(0)}\langle m|H'|n\rangle^{(0)}\,{}^{(0)}\langle n|H'|n\rangle^{(0)}}{\left(E_n^{(0)} - E_m^{(0)}\right)^2} \\ &- \frac{1}{2}\sum_{m\neq n} \frac{|n\rangle^{(0)}\,|\,{}^{(0)}\langle m|H'|n\rangle^{(0)}|^2}{\left(E_n^{(0)} - E_m^{(0)}\right)^2} . \end{aligned} \tag{12.6.51}$$

The use of the second order perturbation formulae is considerably more complicated than the first order perturbations due to the infinite sum over intermediate states. In case there are only a finite number of intermediate states that contribute, the infinite sum reduces to a finite sum. We shall now illustrate this with an example and give a physical interpretation of the perturbation sum.

12.7 Displaced Simple Harmonic Oscillator

The Hamiltonian we want to consider is

$$\begin{aligned} H &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 + \lambda x \\ &= H_0 + \lambda x \ . \end{aligned} \tag{12.7.52}$$

We want to treat λx as a perturbation. Since in terms of creation and annihilation operators

$$x = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger) \ . \tag{12.7.53}$$

We have that to first order in λ

$$\begin{aligned} E_n &= \langle n|H_0 + \lambda x|n\rangle = (n + 1/2)\hbar\omega + \lambda\langle n|x|n\rangle \\ &= (n + 1/2)\hbar\omega \ . \end{aligned} \tag{12.7.54}$$

Then to second order in λ we have

$$\begin{aligned} E_n &= (n + 1/2)\hbar\omega + \lambda^2 \sum_{n\neq r} \frac{\langle n|x|r\rangle\langle r|x|n\rangle}{(n-r)\hbar\omega} \\ &= (n + 1/2)\hbar\omega + \frac{\lambda^2}{2m\omega^2} \sum_{n\neq r} \frac{\langle n|a + a^\dagger|r\rangle\langle r|a + a^\dagger|n\rangle}{(n-r)} \ . \end{aligned} \tag{12.7.55}$$

Using (9.2.60) and (9.2.63) namely,

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{12.7.56}$$

and

$$a|n\rangle = \sqrt{n}|n-1\rangle \tag{12.7.57}$$

we get that to second order in λ

$$E_n = (n + 1/2)\hbar\omega + \frac{\lambda^2}{2m\omega^2} \sum_{n\neq r} \frac{n\delta_{n,r+1} + (n+1)\delta_{n,r-1}}{(n-r)} \tag{12.7.58}$$

or

$$E_n = (n + 1/2)\hbar\omega - \frac{\lambda^2}{2m\omega^2} \ . \tag{12.7.59}$$

If we compare this with the exact solution, which is easily obtained (problem 9.3 with $\beta = 1$ and $V = \lambda\sqrt{\hbar/2m\omega}$) or by completing the square in the Hamiltonian (12.7.52) we get

$$E_n = (n + 1/2)\hbar\omega - \frac{\lambda^2}{2m\omega^2} . \tag{12.7.60}$$

which agrees exactly with the second order perturbation theory. The reason for this is that the exact eigenvalue has only quadratic dependence on λ. Thus, all higher order perturbation terms must vanish. We now try to make plausible, without computation, the absence of third order terms.

In first order perturbation theory the perturbation H' must connect the given unperturbed state $|n\rangle$ with itself if it is to contribute. In second order H' must connect $|n\rangle$ with some other state and then back. Pictorially this is as shown in figure 12.2. This corresponds to the term

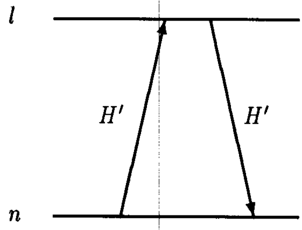

Figure 12.2: Second order perturbation.

$$\frac{\langle n|H'|r\rangle\langle r|H'|n\rangle}{E_n - E_r} \quad r \neq n . \tag{12.7.61}$$

We then sum over all such intermediate states. In the example above we have $r = n \pm 1$ as the only possible intermediate states. In third order the picture is as in figure 12.3. This corresponds to a term

$$\frac{\langle n|H'|m\rangle\langle m|H'|r\rangle\langle r|H'|n\rangle}{(E_n - E_m)(E_m - E_r)} \quad r, m \neq n . \tag{12.7.62}$$

This also explains why there are no third order terms in our previous example. There we could only have either

$$n \to n+1 \begin{matrix} \nearrow & n+2 \\ \searrow & n \end{matrix} \quad \text{(second order)} \tag{12.7.63}$$

or

$$n \to n-1 \begin{matrix} \nearrow & n \\ \searrow & n-2 \end{matrix} \quad \text{(second order)} . \tag{12.7.64}$$

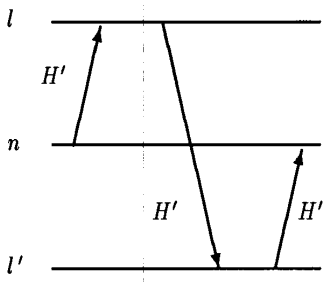

Figure 12.3: Third order perturbation.

In either case there is no non-vanishing matrix element connecting $n \pm 2$ with n. Hence, there are no third order contributions.

We next develop non-degenerate perturbation theory up to arbitrary order. The technique we employ will allow us to recover the Rayleigh-Schrödinger method as well as another method known as the Brillouin-Wigner.

12.8 Non-degenerate Perturbations to all Orders

We again start with the Hamiltonian

$$H = H_0 + \lambda V \tag{12.8.65}$$

and assume that the eigenvalues $E_n^{(0)}$ of H_0 are non-degenerate with eigenkets $|n\rangle^{(0)}$. Suppose we wish to calculate the eigenvalue E_n and corresponding eigenket $|n\rangle$ of H. Thus,

$$(H_0 + \lambda V - E_n)|n\rangle = 0 \ . \tag{12.8.66}$$

We retain the normalization

$$^{(0)}\langle n|n\rangle = 1 \ . \tag{12.8.67}$$

Then using (12.8.66) and (12.8.67) we obtain by using

$$\langle n|H_0|n\rangle^{(0)} = E_n^{(0)} \, \langle n|n\rangle^{(0)} \tag{12.8.68}$$

that

$$E_n^{(0)} + \langle n|V|n\rangle^{(0)} = E_n \ . \tag{12.8.69}$$

This gives us the level shift due to the perturbation.

$$E_n - E_n^{(0)} = \langle n|V|n\rangle^{(0)} = {}^{(0)}\langle n|V|n\rangle \, . \tag{12.8.70}$$

Unfortunately (12.8.70) involves the unknown exact eigenket $|n\rangle$.

To avoid having to exclude the term $m = n$ in the sum over m in the expansion for $|n\rangle$ we define the projection operator (see problem 6.4)

$$P_n = |n\rangle^{(0)}\,{}^{(0)}\langle n| \tag{12.8.71}$$

which has the property of projecting any ket along $|n\rangle^{(0)}$. The operator

$$Q_n = 1 - P_n \tag{12.8.72}$$

is also a projection operator and has the property that any state $Q_n|\psi\rangle$ is always orthogonal to $|n\rangle^{(0)}$. Any ket $|\cdot\rangle$ can now be written as a sum of a term parallel to $|n\rangle^{(0)}$ and a part orthogonal to $|n\rangle^{(0)}$.

$$|\cdot\rangle = P_n|\cdot\rangle + Q_n|\cdot\rangle \ . \tag{12.8.73}$$

Hence, in particular, for the ket $|n\rangle$ we have

$$\begin{aligned} |n\rangle &= P_n|n\rangle + Q_n|n\rangle \\ &= |n\rangle^{(0)} + Q_n|n\rangle \end{aligned} \tag{12.8.74}$$

where we have used (12.8.67).

Now let z be an arbitrary complex number then we may write

$$(H_0 - z)|n\rangle = (E_n - \lambda V - z)|n\rangle \tag{12.8.75}$$

and hence,

$$|n\rangle = \frac{1}{H_0 - z}(E_n - \lambda V - z)|n\rangle \, . \tag{12.8.76}$$

Next, multiply this equation by

$$1 = P_n + Q_n \tag{12.8.77}$$

to get

$$\begin{aligned} |n\rangle &= P_n\frac{1}{H_0 - z}(E_n - \lambda V - z)|n\rangle + Q_n\frac{1}{H_0 - z}(E_n - \lambda V - z)|n\rangle \\ &= P_n|n\rangle + Q_n\frac{1}{H_0 - z}(E_n - \lambda V - z)|n\rangle \ . \end{aligned} \tag{12.8.78}$$

Thus

$$|n\rangle = |n\rangle^{(0)} + Q_n\frac{1}{H_0 - z}(E_n - \lambda V - z)|n\rangle \tag{12.8.79}$$

where in the last step we have used equation (12.8.74).

To get the perturbation series we simply iterate this equation by replacing $|n\rangle$ on the right side in terms of $|n\rangle^{(0)}$. After the first step we get

$$\begin{aligned} |n\rangle &= |n\rangle^{(0)} + Q_n \frac{1}{H_0 - z}(E_n - \lambda V - z)|n\rangle^{(0)} \\ &+ Q_n \frac{1}{H_0 - z}(E_n - \lambda V - z) Q_n \frac{1}{H_0 - z}(E_n - \lambda V - z)|n\rangle \,. \end{aligned} \tag{12.8.80}$$

Repeating this procedure indefinitely we finally obtain

$$|n\rangle = \sum_m \left[Q_n \frac{1}{H_0 - z}(E_n - \lambda V - z) \right]^m |n\rangle^{(0)} \,. \tag{12.8.81}$$

Substituting this in the expression for the energy shift we then get

$$E_n - E_n^{(0)} = \lambda \sum_m {}^{(0)}\langle n| V \left[Q_n \frac{1}{H_0 - z}(E_n - \lambda V - z) \right]^m |n\rangle^{(0)} \,. \tag{12.8.82}$$

By choosing appropriate values for z we now either recover the Rayleigh-Schrödinger theory or else obtain the Brillouin-Wigner Theory.

Thus, choosing $z = E_n^{(0)}$ we recover Rayleigh-Schrödinger perturbation theory.

$$E_n = E_n^{(0)} + \lambda \sum_m {}^{(0)}\langle n| V \left[Q_n \frac{1}{H_0 - E_n^{(0)}}(E_n - \lambda V - E_n^{(0)}) \right]^m |n\rangle^{(0)} \tag{12.8.83}$$

and

$$|n\rangle = \sum_m \left[Q_n \frac{1}{H_0 - E_n^{(0)}}(E_n - \lambda V - E_n^{(0)}) \right]^m |n\rangle^{(0)} \,. \tag{12.8.84}$$

If instead we choose $z = E_n$ we obtain the Brillouin-Wigner perturbation theory. Here,

$$E_n = E_n^{(0)} + \lambda \sum_m {}^{(0)}\langle n| V \left[Q_n \frac{-1}{H_0 - E_n} \lambda V \right]^m |n\rangle^{(0)} \tag{12.8.85}$$

and

$$|n\rangle = \sum_m \left[Q_n \frac{-1}{H_0 - E_n} \lambda V \right]^m |n\rangle^{(0)} \,. \tag{12.8.86}$$

We can use either set to obtain E_n and $|n\rangle$ to arbitrary order in λ. Thus, to first order (12.8.85) yields

$$E_n = E_n^{(0)} + \lambda\, {}^{(0)}\langle n|V|n\rangle^{(0)} \tag{12.8.87}$$

and to second order it yields

$$E_n = E_n^{(0)} + \lambda\, {}^{(0)}\langle n|V|\rangle^{(0)} + \lambda^2 \sum_{m \neq n} \frac{\left| {}^{(0)}\langle n|V|m\rangle^{(0)} \right|^2}{E_n - E_m^{(0)}} \,. \tag{12.8.88}$$

This equation must still be solved for E_n. That is why Rayleigh-Schrödinger perturbation theory is used more often; it gives E_n directly. On the other hand, the Brillouin-Wigner theory converges more rapidly in certain situations and may therefore be more convenient.

These formal expansions to all orders are not any more useful than our previous result in section 12.2 for obtaining numerical values. They are mainly of utility in trying to prove general results.

12.9 Sum Rule: Second Order Perturbation

Second order perturbation theory inevitably gives us sums of the form

$$E_n^{(2)} = \sum_m \frac{{}^{(0)}\langle n|H'|m\rangle^{(0)}\ {}^{(0)}\langle m|H'|n\rangle^{(0)}}{E_n^{(0)} - E_m^{(0)}} \ . \qquad (12.9.89)$$

These sums are infinite in general and cannot be done in closed form. An approximation procedure that has been used on occasion [12.3] is to set $E_m^{(0)} = E$, a constant. Then,

$$\begin{aligned} E_n^{(2)} &= \frac{1}{E_n^{(0)} - E} \sum_{m\neq n} {}^{(0)}\langle n|H'|m\rangle^{(0)}\ {}^{(0)}\langle m|H'|n\rangle^{(0)} \\ &= \frac{\sum_m {}^{(0)}\langle n|H'|m\rangle^{(0)}\ {}^{(0)}\langle m|H'|n\rangle^{(0)} - |{}^{(0)}\langle n|H'|n\rangle^{(0)}|^2}{E_n^{(0)} - E} \ . \end{aligned} \qquad (12.9.90)$$

Using the completeness relation

$$\sum_m |m\rangle^{(0)}\ {}^{(0)}\langle m| = 1 \qquad (12.9.91)$$

the above expression reduces to

$$E_n^{(2)} = \frac{1}{E_n^{(0)} - E}\left[{}^{(0)}\langle n|(H')^2|n\rangle^{(0)} - |{}^{(0)}\langle n|H'|n\rangle^{(0)}|^2\right] \ . \qquad (12.9.92)$$

Thus, we have a neat closed form expression. Unfortunately it still contains the parameter E which must somehow be chosen in an arbitrary fashion. This presents the major drawback for this method.

Fortunately there are quite a few cases in which the sum (12.9.89) can be evaluated exactly if one can solve a certain differential equation. Since one can then do the sum in closed form we obtain a sum rule. In practice there are many different kinds of sum rules [12.4]; we concentrate here only on the particular sum expressed by equation (12.9.89).

To evaluate this perturbation sum we look for an operator F_n such that

$$H'|n\rangle^{(0)} = [F_n, H_0]|n\rangle^{(0)} \ . \qquad (12.9.93)$$

Here as before

$$H = H_0 + \lambda H' \ . \qquad (12.9.94)$$

Since the right hand side of (12.9.93) vanishes if we take an inner product with ${}^{(0)}\langle n|$ we have to modify this equation slightly to read

$$ {}^{(0)}\langle m|F_n|n\rangle^{(0)} = E_n^1\delta_{n,m} + {}^{(0)}\langle m|[F_n, H_0]|n\rangle^{(0)} . \tag{12.9.95}$$

This expression clearly simplifies since

$$\begin{aligned} {}^{(0)}\langle m|[F_n, H_0]|n\rangle^{(0)} &= {}^{(0)}\langle m|F_nH_0 - H_0F_n|n\rangle^{(0)} \\ &= {}^{(0)}\langle m|F_nE_n^{(0)} - E_m^{(0)}F_n|n\rangle^{(0)} \\ &= \left(E_n^{(0)} - E_m^{(0)}\right) {}^{(0)}\langle m|F_n|n\rangle^{(0)} . \end{aligned} \tag{12.9.96}$$

Substituting this result on the right side of (12.6.48) we find

$$\begin{aligned} E_n^{(2)} &= \sum_{m\neq n} {}^{(0)}\langle n|H'|m\rangle^{(0)}\ {}^{(0)}\langle m|F_n|n\rangle^{(0)} \\ &= \sum_{m} {}^{(0)}\langle n|H'|m\rangle^{(0)}\ {}^{(0)}\langle m|F_n|n\rangle^{(0)} \\ &- \ {}^{(0)}\langle n|H'|n\rangle^{(0)}\ {}^{(0)}\langle n|F_n|n\rangle^{(0)} \end{aligned} \tag{12.9.97}$$

or

$$ E_n^{(2)} ={}^{(0)} \langle n|H'F_n|n\rangle^{(0)} - E_n^{(1)}\ {}^{(0)}\langle n|F_n|n\rangle^{(0)} . \tag{12.9.98}$$

For the last step we have again used the completeness relation

$$ \sum_m |m\rangle^{(0)}\ {}^{(0)}\langle m| = 1 \tag{12.9.99}$$

as well as the fact that

$$ E_n^{(1)} = {}^{(0)}\langle n|H'|n\rangle^{(0)} . \tag{12.9.100}$$

Thus, if we can find such an F_n then the sum for the energy in second order perturbation theory can be reduced to an integral. To see how this F_n may be evaluated in practice, consider the following Hamiltonian for a system in one dimension

$$ H = \frac{p^2}{2m} + V(x) + \lambda U(x) \tag{12.9.101}$$

where the unperturbed Hamiltonian

$$ H = \frac{p^2}{2m} + V(x) \tag{12.9.102}$$

has the eigenfunctions $\phi_n(x)$ corresponding to the non- degenerate eigenvalues $E_n^{(0)}$. We now assume that the operator F_n is simply a function of x. Equation (12.9.93) when written out then reads:

$$\begin{aligned} U(x)\phi_n(x) &= -\frac{\hbar^2}{2m}\left\{F_n(x)\frac{d^2\phi_n}{dx^2} - \frac{d^2}{dx^2}\left[F_n(x)\phi_n(x)\right]\right\} \\ &= \frac{\hbar^2}{2m}\left\{\frac{d^2F_n}{dx^2}\phi_n(x) + 2\frac{dF_n(x)}{dx}\frac{d\phi_n(x)}{dx}\right\} . \end{aligned} \tag{12.9.103}$$

Thus,

$$\frac{d^2F_n}{dx^2}+2\frac{\phi_n'(x)}{\phi_n(x)}\frac{dF_n(x)}{dx}=\frac{2m}{\hbar^2}U(x)\,. \tag{12.9.104}$$

Calling

$$\frac{dF_n(x)}{dx}=f(x) \tag{12.9.105}$$

we get the first order linear equation

$$\frac{df}{dx}+2\frac{\phi_n'}{\phi_n}f=\frac{2m}{\hbar^2}U(x) \tag{12.9.106}$$

which may be integrated by multiplying by the integrating factor

$$R=\phi_n^2(x) \tag{12.9.107}$$

to give

$$\frac{d}{dx}\left(\phi_n^2 f\right)=\frac{2m}{\hbar^2}\phi_n^2(x)U(x)\,. \tag{12.9.108}$$

Hence, up to an integration constant

$$f(x)\equiv\frac{dF_n}{dx}=\frac{2m}{\hbar^2}\frac{1}{\phi_n^2}\int_a^x U(y)\phi_n^2(y)\,dy\ . \tag{12.9.109}$$

Integrating once more, we find

$$F_n(x)=\frac{2m}{\hbar^2}\int_b^x\frac{dy}{\phi_n^2(y)}\int_a^y U(z)\phi_n^2(z)\,dz\,. \tag{12.9.110}$$

The constants a and b simply determine the value of $F_n(b)=0$ and $dF_n/dx\,|_{x=a}=0$. Since these values can always be changed by adding to our *particular solution* (12.9.110) an arbitrary solution of the homogeneous equation

$$\frac{d^2F_n}{dx^2}+2\frac{\phi_n'(x)}{\phi_n(x)}\frac{dF_n(x)}{dx}=0 \tag{12.9.111}$$

they have no effect on the sum we want to evaluate. The reason for this is that solutions of the homogeneous equation correspond to $U(x)=0$. This simply means that, as far as our perturbation sum is concerned, the boundary conditions on F_n are arbitrary and may be chosen for convenience as above.

We have now succeeded in reducing the sum for the second order energy perturbation to integrals over known functions. In practice these integrals can seldom be done in closed form. However, whenever the perturbation sum is infinite, the integrals are easier to evaluate numerically than the sum.

12.10 Linear Stark Effect

A practical example using the sum rules just developed, was first worked out by Dalgarno and Lewis [12.6] and yields, to second order, the shift in the ground state energy of an hydrogen atom due to a constant electric field. This is known as the linear Stark effect. In this case the unperturbed Hamiltonian H_0 is just the hydrogen Hamiltonian

$$H_0 = \frac{\mathbf{p}^2}{2m} - \frac{e^2}{r} . \tag{12.10.112}$$

The corresponding ground state wavefunction is

$$\phi_0 = \frac{1}{\sqrt{\pi a^3}} e^{-r/a} . \tag{12.10.113}$$

The perturbation term is

$$H' = e\mathcal{E}z \tag{12.10.114}$$

where $\mathcal{E}$ is the strength of the electric field.

This means that the operator F_0 must satisfy the equation

$$e\mathcal{E}ze^{-r/a} = [F_0, H_0]e^{-r/a} . \tag{12.10.115}$$

We try to solve this equation by choosing

$$F_0 = zf(r) = \cos\theta\, r f(r) . \tag{12.10.116}$$

Straightforward algebra then yields the following differential equation for $f(r)$

$$r\frac{d^2 f}{dr^2} + \left(4 - \frac{2r}{a}\right)\frac{df}{dr} - \frac{2}{a}f = 2\alpha r \tag{12.10.117}$$

where

$$\alpha = \frac{m e\mathcal{E}}{\hbar^2} . \tag{12.10.118}$$

A particular integral of this equation is

$$f(r) = -\alpha a(r/2 + a) . \tag{12.10.119}$$

Thus,

$$F_0 = -\alpha a z(r/2 + a) . \tag{12.10.120}$$

It is now a simple matter to evaluate the energy shift

$$E_0^{(2)} = \frac{-1}{\pi a^3} \int e^{-2r/a} \frac{\alpha^2\hbar^2}{m} a z^2 (r/2 + a) r^2\, dr \sin\theta d\theta d\varphi . \tag{12.10.121}$$

Since the directions x, y, z are all equivalent we can replace z^2 by

$$\frac{1}{3}(x^2 + y^2 + z^2) = \frac{1}{3}r^2 .$$

The θ and φ integrations can then be performed immediately and we arrive at

$$E_0^{(2)} = -\frac{4}{a^2}\frac{\alpha^2\hbar^2}{3m}\int_0^\infty e^{-2r/a}(r/2+a)r^4\,dr\ . \tag{12.10.122}$$

This integral is easily evaluated and we find

$$E_0^{(2)} = -\frac{9}{4}\frac{\alpha^2\hbar^2}{3m} = -\frac{9}{4}a^3\mathcal{E}^2\ . \tag{12.10.123}$$

This result is the same as what we would obtain if we were to evaluate the following sum plus integral

$$-(e\mathcal{E})^2\left[\sum_{n\neq 1,l,m}\frac{|\langle 1,0,0|z|n,l,m\rangle|^2}{\frac{e^2}{2a_0}\left(1-1/n^2\right)}+\int d^3k\frac{|\langle 1,0,0|z|\mathbf{k}\rangle|^2}{\frac{e^2}{2a_0}+\frac{\hbar^2\mathbf{k}^2}{2m}}\right] \tag{12.10.124}$$

where $|n,l,m\rangle$ and $|\mathbf{k}\rangle$ are hydrogen atom eigenkets.

We have treated the effect of a constant electric field as a perturbation in spite of the fact that for $z\to\infty$ the total potential energy due to the perturbation tends to $-\infty$. This means that an exact solution of the perturbed Schrödinger equation would no longer yield bound states, but would lead to a continuous spectrum. The total potential as a function of z (for fixed x and y) looks as shown in figure 12.4. This picture also yields the physical reason why the perturbative solution makes sense. An electron trapped in the region between z_1 and z_2 has a finite, but exceedingly small probability to tunnel through to the region $z > z_3$ where it would become unbound. The probability for this to happen is, however, so low that the quasi-bound state that we used in our calculation still makes sense and gives useful results.

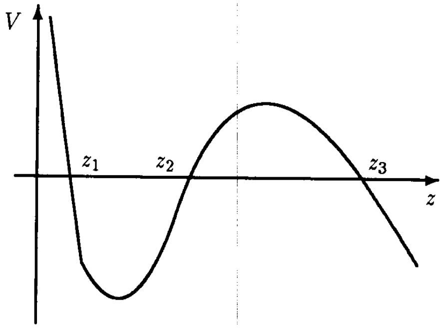

Figure 12.4: The Stark effect.

This completes our treatment of non-degenerate perturbation theory. In the next chapter we develop perturbation theory for states which are degenerate in energy.

12.11 Problems

12.1 Verify formula (12.5.42).
Hint: Solve the problem

$$\nabla^2\psi(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}')$$

by
a) expanding in spherical harmonics,
b) realizing that $\psi(\mathbf{r})$ is the potential for a unit charge located at $\mathbf{r}'$, and comparing the two solutions.

12.2 In (12.2.17) set $r = 2$ and derive equation (12.6.51) for the second order correction to the wavefunction.

12.3 Consider the Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 + \frac{1}{2}\lambda x^2 \quad k > 0 \ .$$

a) Find the exact energy of the n'th state of this Hamiltonian and expand it to order λ^2 assuming $|\lambda| < k$.
b) Use perturbation theory, treating $(1/2)\lambda x^2$ as a perturbation, and find the energy of the nth state to order λ^2 .
c) Find the rth order correction and hence show that the perturbation series converges for $|\lambda| < k$.

12.4 a) Find the approximate ground state energy to second order for the Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 + \frac{1}{4}\lambda x^4 \quad k > 0$$

using the Rayleigh-Schrödinger perturbation theory.
b) Find the ground state correct to order λ.

12.5 Repeat problem (12.4a) using Brillouin-Wigner perturbation theory.

12.6 Consider the Hamiltonian

$$H = H_0 + \lambda H'$$

where

$$H_0 = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \quad H' = \begin{pmatrix} 0 & ia \\ -ia & 0 \end{pmatrix} \ .$$

a) Solve for the exact eigenvalues and eigenfunctions.
b) Solve for both eigenvalues and eigenfunctions to 2nd order using Rayleigh-Schrödinger perturbation theory.

12.7 A particle of mass m moves in a potential

$$V = \frac{1}{2}k|x|^{2+\epsilon} \qquad |\epsilon| < 1 \, .$$

Estimate the energy of the ground state.
Hint:

$$\frac{1}{2}k|x|^{2+\epsilon} = \frac{1}{2}kx^2 - \frac{1}{2}k(x^2 - |x|^{2+\epsilon}) \approx \frac{1}{2}kx^2 + \frac{\epsilon}{2}kx^2 \ln|x| \, .$$

Also,

$$\int_0^{\infty} e^{-\alpha x^2} x^2 \ln x \, dx = \frac{\sqrt{\pi}}{4} a^{-3/2} \left[1 - \frac{1}{2}(c + \ln 4a)\right]$$

where

$$c = 0.577216\ldots = \text{Euler's constant} \, .$$

12.8 For a particle of mass m moving in the potential

$$V = \frac{1}{2}k_1 x^2 + \frac{2}{2}k_1 x^2 + \lambda xy \qquad |k_1 - k_2| > 2\lambda$$

a) find the exact energy levels.
b) Use perturbation theory to find to order λ^2 the energy of all the levels and compare with the exact solution to this order.

12.9 Kuhn-Thomas-Reiche Sum Rule
Classically the polarizability α of an atom is defined as the induced electric dipole moment $e|\vec{r}|$ divided by the strength of the inducing electric field $\vec{E}$. So,

$$\alpha = \frac{e|\vec{r}|}{|\vec{E}|}$$

and for harmonically bound electrons takes the form

$$\alpha = \frac{e^2}{4\pi^2 m} \sum_j \frac{f_j}{\nu_j^2 - \nu^2} \quad .$$

Here, f_j are dimensionless constants called the "oscillator strengths". In quantum mechanics these are defined by

$$\begin{aligned} f_j &= \frac{4\pi m}{3\hbar e^2} \nu_{j0} |\vec{m}_{j0}|^2 \\ &= \frac{4\pi m}{3\hbar e^2} \frac{E_j - E_0}{h} |e\vec{r}_{j0}|^2 \\ &= \frac{2m}{3\hbar^2} (E_j - E_0) |\vec{r}_{j0}|^2 \end{aligned}$$

where

$$|\vec{m}_{j0}| = \langle|\vec{m}_{j0}|0\rangle$$

and

$$|\vec{r}_{j0}| = \langle|\vec{r}_{j0}|0\rangle \ .$$

For N uncoupled electrons one then has the Kuhn-Thomas-Reiche sum rule

$$\sum_j f_j = N \ .$$

This polarizability can be used to describe the absorption of light which carries an electron from its ground state $|0\rangle$ to an excited state $|n\rangle$ in an atom. If, the Hamiltonian for the bound electron is

$$H = \frac{\vec{p}^2}{2m} + V(\vec{r}) \ .$$

a) Show that

$$(E_n - E_0)\vec{r}_{n0} = -\frac{i\hbar}{m}\langle n|\vec{p}|0\rangle \ .$$

Hint: Use the commutator $[H, \vec{r}]$ and work component by component.
b) Use the commutators

$$[x, p_x] = [y, p_y] = [z, p_z] = i\hbar$$

together with the results of part a) to prove that for a single electron

$$\sum_n \frac{2m(E_n - E_0)}{3\hbar^2}[|x_{n0}|^2 + |y_{n0}|^2 + |z_{n0}|^2] = 1$$

and hence deduce the Kuhn-Thomas-Reiche sum rule.

Bibliography

[12.1] One of the original references on perturbation theory and still very readable today is
J.W.S. Rayleigh, *The Theory of Sound - Vol. I* sect. 90, Dover Publications, New York (1945).

[12.2] A very concise treatment of Rayleigh-Schrödinger perturbation theory is given in
E.P. Wigner, *Group Theory* - Chapter 5, Academic Press, New York (1959).

[12.3] A. Unsold, Z. Phys. **43**, 563 (1967).

[12.4] R. Jackiw, Quantum-mechanical sum rules - Phys. Rev. **157**, 1220-1225 (1967) item[12.5] C. Schwartz, Calculations in Schrödinger Perturbation theory - Annals of Physics N.Y. **2**, 156-169, (1959).

[12.6] A Dalgarno and J.T. Lewis, The exact calculation of long-range forces between atoms by perturbation - Proc. Roy. Soc. **A233**, 70-84, (1956).

Chapter 13

Degenerate Perturbation Theory

13.1 Introduction

In many cases of physical interest the energy levels of a simple Hamiltonian are degenerate. This is always the case if the simple Hamiltonian is invariant under some symmetry operation such as a spatial rotation or reflection. Thus, for example all levels except the ground state of the three-dimensional simple harmonic oscillator show the $2l+1$ fold degeneracy arising from rotation symmetry. A similar statement is true for the hydrogen atom although in this case the degeneracy is larger since an even larger symmetry group than just rotations leaves the hydrogen Hamiltonian invariant.

Now in order that the perturbation expansions (12.2.6), (12.2.7) be meaningful and useful, it is necessary that the series converge relatively fast. In fact to be useful we should be able to truncate the series after the first few terms for otherwise the computations become excessive. This further requires that the successive terms in the series decrease rapidly. On the other hand if we have degenerate or even just close-lying energy levels, the second order terms

$$\lambda^2 \sum_{m \neq n} \frac{|H'_{m,n}|^2}{E_n^{(0)} - E_m^{(0)}}$$

will give an excessively large contribution whenever the m and n levels are degenerate or close. This is due to our treatment and we now develop techniques to handle this situation.

13.2 Two Levels: Rayleigh-Schrödinger Method

In this section we consider the case of only two levels, say $|1,1\rangle^{(0)}$ and $|1,2\rangle^{(0)}$ that are degenerate. As discussed above, for perturbation theory to be applicable at all, we require that

$$\lambda|H'_{n,m}| < |E_n^{(0)} - E_m^{(0)}| \quad n \neq m \ . \tag{13.2.1}$$

Our Hamiltonian, as always, splits into

$$H = H_0 + \lambda H' \tag{13.2.2}$$

with the unperturbed, nondegenerate states given by

$$H_0|n\rangle^{(0)} = E_n^{(0)}|n\rangle^{(0)} \quad n \neq 1 \tag{13.2.3}$$

and the unperturbed, degenerate states given by

$$\begin{aligned} H_0|1,1\rangle^{(0)} &= E_1^{(0)}|1,1\rangle^{(0)} \\ H_0|1,2\rangle^{(0)} &= E_1^{(0)}|1,2\rangle^{(0)} \ . \end{aligned} \tag{13.2.4}$$

Here we have explicitly shown that two eigenvalues say $E_1^{(0)}$ of H_0 are degenerate. In this case condition (13.2.1) will be violated for the two levels $|1,1\rangle^{(0)}$ and $|1,2\rangle^{(0)}$. Suppose further that all the other energy levels that are non-degenerate are widely spaced. In that case the sum

$$\lambda^2 E_n^{(2)} = \lambda^2 \sum_{m\neq n} \frac{|H'_{n,m}|^2}{E_n^{(0)} - E_m^{(0)}} \tag{13.2.5}$$

diverges for $m = 1$ unless the matrix element

$$H'_{1,2} = {H'_{2,1}}^* = {}^{(0)}\langle 1,1|H'|1,2\rangle^{(0)} \tag{13.2.6}$$

vanishes. This vanishing can indeed be accomplished due to the fact that, in the subspace spanned by the eigenfunctions corresponding to the degenerate eigenvalues, any linear combination of the eigenfunctions is also an eigenfunction. In other words, there is no reason to prefer the basis which starts with $|1,1\rangle^{(0)}$, $|1,2\rangle^{(0)}$ over any other basis set which starts with a linear combination of $|1,1\rangle^{(0)}$ and $|1,2\rangle^{(0)}$ and the remaining basis kets $|2\rangle^{(0)}$, $|3\rangle^{(0)}$, ... etc. Thus, instead of the basis set

$$|1,1\rangle^{(0)} \ , \ |1,2\rangle^{(0)} \ , \ |2\rangle^{(0)} \ , \ |3\rangle^{(0)} \ , \ |4\rangle^{(0)} \ , \ \ldots$$

we can start with the set

$$a_{11}|1,1\rangle^{(0)} + a_{12}|1,2\rangle^{(0)} \ , \ a_{21}|1,1\rangle^{(0)} + a_{22}|1,2\rangle^{(0)} \ , \ |2\rangle^{(0)} \ , \ |3\rangle^{(0)} \ , \ |4\rangle^{(0)} \ , \ \ldots$$

where the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{13.2.7}$$

is unitary. We can now further specify the matrix A by requiring that it diagonalize H' in the degenerate subspace spanned by $|1,1\rangle^{(0)}$ and $|1,2\rangle^{(0)}$. So, defining the matrix

$$V = \begin{pmatrix} H'_{11} & H'_{12} \\ H'_{21} & H'_{22} \end{pmatrix} \tag{13.2.8}$$

where

$$H'_{\mu,\nu} = {}^{(0)}\langle 1,\mu|H'|1,\nu\rangle^{(0)} \quad \mu = 1,2 \quad \nu = 1,2 \ . \tag{13.2.9}$$

We require that $AVA^\dagger$ be diagonal. This is accomplished by solving the following eigenvalue problem

$$V\begin{pmatrix} a \\ b \end{pmatrix} = v\begin{pmatrix} a \\ b \end{pmatrix} \tag{13.2.10}$$

and forming the matrix A from the eigenvectors. For a non-trivial solution we require

$$\det(V - v1) = 0 \tag{13.2.11}$$

giving us the characteristic equation

$$(H'_{1,1} - v)(H'_{2,2} - v) - |H'_{1,2}|^2 = 0 \ . \tag{13.2.12}$$

The roots of this equation are

$$v_\pm = \frac{H'_{11} + H'_{22}}{2} \pm \frac{1}{2}\sqrt{(H'_{11} - H'_{22})^2 + 4|H'_{1,2}|^2} \ . \tag{13.2.13}$$

If we now define

$$\tan\theta = \frac{2|H'_{1,2}|}{H'_{11} - H'_{22}} \tag{13.2.14}$$

we get that the corresponding eigenvectors satisfy

$$\begin{aligned} \frac{a_+}{b_+} &= \cot\frac{\theta}{2} \\ \frac{a_-}{b_-} &= -\tan\frac{\theta}{2} \end{aligned} \tag{13.2.15}$$

so that the normalized eigenvectors are

$$\begin{pmatrix} a_+ \\ b_+ \end{pmatrix} = \begin{pmatrix} \cos\theta/2 \\ \sin\theta/2 \end{pmatrix} \quad , \quad \begin{pmatrix} a_- \\ b_- \end{pmatrix} = \begin{pmatrix} -\sin\theta/2 \\ \cos\theta/2 \end{pmatrix} \tag{13.2.16}$$

and we get:

$$\begin{aligned} |\psi_{1,1}\rangle &= \cos\frac{\theta}{2}|1,1\rangle^{(0)} + \sin\frac{\theta}{2}|1,2\rangle^{(0)} \\ |\psi_{1,2}\rangle &= -\sin\frac{\theta}{2}|1,1\rangle^{(0)} + \cos\frac{\theta}{2}|1,2\rangle^{(0)} \ . \end{aligned} \tag{13.2.17}$$

Therefore, we now use $|\psi_{1,1}\rangle$, $|\psi_{1,2}\rangle$ as the first two members of our basis. This ensures that in the degenerate subspace H' is diagonal so that all its off-diagonal elements vanish. This means that no correction appears for these levels in first order and in the perturbation to second order the two divergent terms also do not appear. We make all of this more explicit now.

Having settled on the basis states with which we want to start, we proceed with standard perturbation theory as developed in the previous chapter. However, as stated, we use the set

$$\{|\psi\rangle\} = \{|\psi_{1,1}\rangle\ ,\ |\psi_{1,2}\rangle\ ,\ |2\rangle^{(0)}\ ,\ |3\rangle^{(0)}\ ,\ |4\rangle^{(0)}\ ,\ \ldots\}$$

as a basis set from the start. Thus, as always, we let

$$E_{1,\nu} = E_1^{(0)} + \lambda E_{1,\nu}^{(1)} + \lambda^2 E_{1,\nu}^{(2)} + \ldots \quad (\nu = 1,2)\ . \tag{13.2.18}$$

Also the exact wavefunctions $|1,\nu\rangle$ (to order λ^2) for ($\nu = 1,2$) are given by

$$\begin{aligned} |1,1\rangle &= |\psi_{1,1}\rangle + \lambda b_{1,2}^{(1)}|\psi_{1,2}\rangle + \lambda \sum_{k\neq 1} a_{1,k}^{(1)}|k\rangle^{(0)} \\ &+ \lambda^2 b_{1,2}^{(2)}|\psi_{1,2}\rangle + \lambda^2 \sum_{k\neq 1} a_{1,k}^{(2)}|k\rangle^{(0)} \\ &= |\psi_{1,1}\rangle + \lambda|1,1\rangle^{(1)} + \lambda^2|1,1\rangle^{(2)} \end{aligned} \tag{13.2.19}$$

$$\begin{aligned} |1,2\rangle &= |\psi_{1,2}\rangle + \lambda b_{2,1}^{(1)}|\psi_{1,1}\rangle + \lambda \sum_{k\neq 1} a_{2,k}^{(1)}|k\rangle^{(0)} \\ &+ \lambda^2 b_{2,1}^{(2)}|\psi_{1,1}\rangle + \lambda^2 \sum_{k\neq 1} a_{2,k}^{(2)}|k\rangle^{(0)} \\ &= |\psi_{1,2}\rangle + \lambda|1,2\rangle^{(1)} + \lambda^2|1,2\rangle^{(2)}\ . \end{aligned} \tag{13.2.20}$$

Substituting these expansions into the Schrödinger equation

$$(H_0 + \lambda H')|1,\nu\rangle = E_{1,\nu}|1,\nu\rangle \tag{13.2.21}$$

and proceeding to second order in λ we arrive (for $\nu = 1,2$) at

$$H_0|\psi_{1,\nu}\rangle = E_1^{(0)}|\psi_{1,\nu}\rangle \tag{13.2.22}$$

$$(H_0 - E_1^{(0)})|1,\nu\rangle^{(1)} + (H' - E_{1,\nu}^{(1)})|\psi_{1,\nu}\rangle = 0 \tag{13.2.23}$$

$$(H_0 - E_1^{(0)})|1,\nu\rangle^{(2)} + (H' - E_{1,\nu}^{(1)})|1,\nu\rangle^{(1)} - E_{1,\nu}^{(2)}|\psi_{1,\nu}\rangle = 0\ . \tag{13.2.24}$$

These equations may now be solved just as in the non-degenerate case. Writing them out, using (13.2.19) and (13.2.20), we get to first order

$$\sum_{k\neq 1} a_{1,k}^{(1)}(E_k^{(0)} - E_1^{(0)})|k\rangle^{(0)} + (H' - E_{1,\nu}^{(1)})|\psi_{1,\nu}\rangle = 0 \tag{13.2.25}$$

where we have already used the fact that the energies for the states $|\psi_{1,\nu}\rangle$ ($\nu = 1,2$) are degenerate. To second order we find

$$\begin{aligned} &\sum_{k\neq 1} a_{1,k}^{(2)}(E_k^{(0)} - E_1^{(0)})|k\rangle^{(0)} \\ &+ \ (H' - E_{1,\nu}^{(1)})\left[b_{\nu,\mu}^{1}|\psi_{1,\mu}\rangle + \sum_{k\neq 1} a_{1,k}^{(1)}|k\rangle^{(0)}\right] - E_{1,\nu}^{(2)}|\psi_{1,\nu}\rangle \\ &= \ 0\ . \end{aligned} \tag{13.2.26}$$

Solving in the usual manner, by taking inner products, first with $\langle\psi_{1,\nu}|$ we find from (13.2.25) that

$$E^{(1)}_{1,\nu} = \langle\psi_{1,\nu}|H'|\psi_{1,\nu}\rangle = v_\pm \tag{13.2.27}$$

and then taking the inner product with ${}^{(0)}\langle k|$

$$a^{(1)}_{\nu,k} = \frac{{}^{(0)}\langle k|H'|\psi_{1,\nu}\rangle}{E^{(0)}_1 - E^{(0)}_k} \qquad k \neq 1\ . \tag{13.2.28}$$

However, we do not get any equations for $b^{(1)}_{\nu,\mu}$ from the first order equations (13.2.25). In fact, the equations for these coefficients come from the second order equations (13.2.26). At this stage we just write down the answer since we repeat the calculation, in detail, for the general case in the next section.

$$\begin{aligned} |1,\nu\rangle^{(1)} &= \sum_{m\neq 1}\left\{\frac{|\psi_{1,\mu}\rangle\langle\psi_{1,\mu}|H'|m\rangle^{(0)}\,{}^{(0)}\langle m|H'|\psi_{1,\nu}\rangle}{(E^{(0)}_1 - E^{(0)}_m)(E^{(1)}_{1,\nu} - E^{(1)}_{1,\mu})}\right. \\ &+ \left.\frac{|m\rangle^{(0)}\,{}^{(0)}\langle m|H'|\psi_{1,\nu}\rangle}{E^{(0)}_1 - E^{(0)}_m}\right\} \\ &\text{for} \quad \mu \neq \nu\ . \end{aligned} \tag{13.2.29}$$

The states that are not degenerate are perturbed in the same manner as before, namely:

$$|n\rangle = |n\rangle^{(0)} + \lambda\sum_{m\neq n}\frac{|m\rangle^{(0)}\,{}^{(0)}\langle m|H'|n\rangle}{E^{(0)}_n - E^{(0)}_m}\ . \tag{13.2.30}$$

The second order term for the energy is then given by:

$$E^{(2)}_{1,\nu} = \sum_{m=2}^{\infty}\frac{\left|\langle\psi_{1,\nu}|H'|m\rangle^{(0)}\right|^2}{E^{(0)}_1 - E^{(0)}_m} \qquad \nu = 1,2 \tag{13.2.31}$$

and

$$E^{(2)}_n = \sum_{m\neq n}^{\infty}\frac{\left|H'_{n,m}\right|^2}{E^{(0)}_n - E^{(0)}_m} \qquad n \neq 1\ . \tag{13.2.32}$$

This technique will, of course, only work if the perturbation H' removes the degeneracy in first order, that is, if $v_\pm$ are distinct and sufficiently widely separated. Otherwise one must perform an exact diagonalization of the matrix corresponding to second order in the degenerate subspace. The case discussed is the most important in practice and we do not discuss the more complicated situation where one must go to second order to remove the degeneracy.

13.3 Rayleigh-Schrödinger: Degenerate Levels

We now repeat the calculation of the previous section, but this time we assume that every energy eigenvalue $E_n^{(0)}$ of the unperturbed Hamiltonian is g_n-fold degenerate. Again the Hamiltonian is

$$H = H_0 + \lambda H' \ . \tag{13.3.33}$$

The eigenfunctions corresponding to the degenerate eigenvalue $E_n^{(0)}$ are

$$|n,1\rangle^{(0)} \ , \ |n,2\rangle^{(0)} \ , \ \dots \ , \ |n,g_n\rangle^{(0)}.$$

As discussed in the previous section, the difficulty due to degeneracy which is obvious in second order perturbation for the energy will not occur if the degeneracy can be lifted in first order by diagonalizing H' in each of the degenerate subspaces. To do this we take a new basis set $\{|\psi_{m,\mu}\rangle \quad \mu = 1,2,\dots,g_m\}$ obtained from the original set $\{|n,\nu\rangle^{(0)}\}$ by a unitary transformation. Thus,

$$|\psi_{m,\mu}\rangle = \sum_{\mu'=1}^{g_m} a^m_{\mu,\mu'} |m,\mu'\rangle^{(0)} \ . \tag{13.3.34}$$

We further choose the set $\{|\psi_{m,\mu}\rangle\}$ so that using it as a basis diagonalizes H' in the corresponding g_m-dimensional subspace of interest. We therefore have to solve the following set of eigenvalue equations

$$\langle\psi_{m,\mu}|H'|\psi_{m,\mu'}\rangle = E^{(1)}_{m,\mu}\delta_{\mu,\mu'} \tag{13.3.35}$$

where $E^{(1)}_{m,\mu}$ are the resultant eigenvalues. This then reads

$$\sum_{\mu'=1}^{g_m} \left[{}^{(0)}\langle m,\mu|H'|m,\mu'\rangle^{(0)} - E^{(1)}_{m,\mu}\delta_{\mu,\mu'}\right] a^m_{\mu',\kappa} = 0 \ . \tag{13.3.36}$$

Here μ is a label to distinguish the different eigenvalues that were degenerate in lowest order.

Assuming that all the eigenvalues of (13.3.36) are distinct and well separated, we now proceed with the usual perturbation theory using the eigenfunctions $|\psi_{m,\mu}\rangle$ from the start. Thus, we let

$$E_{m,\mu} = E_m^{(0)} + \lambda E^{(1)}_{m,\mu} + \lambda^2 E^{(2)}_{m,\mu} + \dots \tag{13.3.37}$$

and

$$|m,\mu\rangle = |\psi_{m,\mu}\rangle + \lambda|m,\mu\rangle^{(1)} + \lambda^2|m,\mu\rangle^{(2)} + \dots \ . \tag{13.3.38}$$

Substituting these expansions into the Schrödinger equation

$$H|m,\mu\rangle = E_{m,\mu}|m,\mu\rangle \tag{13.3.39}$$

we obtain the equations corresponding to (12.2.9) to (12.2.12). Thus, we have

$$H_0|\psi_{m,\mu}\rangle = E_m^{(0)}|\psi_{m,\mu}\rangle \tag{13.3.40}$$

$$(H_0 - E_m^{(0)})|m,\mu\rangle^{(1)} + (H' - E_{m,\mu}^{(1)})|\psi_{m,\mu}\rangle = 0 \tag{13.3.41}$$

$$(H_0 - E_m^{(0)})|m,\mu\rangle^{(2)} + (H' - E_{m,\mu}^{(1)})|m,\mu\rangle^{(1)} - E_{m,\mu}^{(2)}|\psi_{m,\mu}\rangle = 0 \ . \tag{13.3.42}$$

Just like the case of only two degenerate levels, we again expand the state $|m,\mu\rangle$ in terms of the unperturbed eigenkets $|\psi_{m,\mu'}\rangle$ $\mu' \neq \mu$ and the states $|k,\kappa\rangle^{(0)}$ $k \neq m$. Then,

$$|m,\mu\rangle^{(1)} = \sum_{\mu'\neq\mu}^{g_m} b_{\mu,\mu'}^{(1)}|\psi_{m,\mu'}\rangle + \sum_{k\neq m}\sum_{\kappa=1}^{g_k} a_{m\mu,k\kappa}^{(1)}|k,\kappa\rangle^{(0)} \tag{13.3.43}$$

with a similar expression for $|m,\mu\rangle^{(2)}$. The first order perturbations are now obtained from (13.3.41). So, by taking the inner product with $\langle\psi_{m,\mu}|$ we find

$$E_{m,\mu}^{(1)} = \langle\psi_{m,\mu}|H'|\psi_{m,\mu}\rangle \ . \tag{13.3.44}$$

Also the coefficients $a_{m\mu,k\kappa}^{(1)}$ that give part of the first order corrections to the wave function are obtained from (13.3.41) by simply taking the inner product of that equation with the bra ${}^{(0)}\langle k,\kappa|$.

To obtain the first order coefficients $b_{\mu,\mu'}^{(1)}$ we have to use the second order equations (13.3.42). These read in analogy to equations (13.2.26)

$$\begin{aligned}
& \sum_{k\neq m}\sum_{\kappa=1}^{g_k} a_{m\mu,k\kappa}^{(2)}[E_k^{(0)} - E_m^{(0)}]|k,\kappa\rangle^{(0)} \\
+ \quad & (H' - E_{m,\mu}^{(1)})\left[\sum_{\mu'\neq\mu}^{g_m} b_{\mu,\mu'}^{(1)}|\psi_{m,\mu'}\rangle + \sum_{k\neq m}\sum_{\kappa=1}^{g_k} a_{m\mu,k\kappa}^{(1)}|k,\kappa\rangle^{(0)}\right] \\
= \quad & E_{m,\mu}^{(2)}|\psi_{m,\mu}\rangle \ .
\end{aligned} \tag{13.3.45}$$

Taking the inner product with the bra $\langle\psi_{m,\mu'}|$ $\mu' \neq \mu$ we find

$$(E_{m,\mu'}^{(1)} - E_{m,\mu}^{(1)})b_{\mu,\mu'}^{(1)} + \sum_{k\neq m}\sum_{\kappa=1}^{g_k} a_{m\mu,k\kappa}^{(1)}\langle\psi_{\mu,\mu'}|H'|k,\kappa\rangle^{(0)} = 0 \ . \tag{13.3.46}$$

Thus, we have solved for $b_{\mu,\mu'}^{(1)}$ and we have the wave function correct to first order.

$$\begin{aligned}
|m,\mu\rangle^{(1)} \quad = \quad & \sum_{k\neq m}\sum_{\kappa=1}^{g_m}\sum_{\mu'\neq\mu} \frac{|\psi_{m,\mu'}\rangle\langle\psi_{m,\mu'}|H'|k,\kappa\rangle^{(0)} \ {}^{(0)}\langle|H'|\psi_{m,\mu}\rangle}{(E_{m,\mu}^{(1)} - E_{m,\mu'}^{(1)})(E_m^{(0)} - E_k^{(0)})} \\
+ \quad & \sum_{k\neq m}\sum_{\kappa=1}^{g_m} \frac{|k,\kappa\rangle^{(0} \ {}^{(0)}\langle|H'|\psi_{m,\mu}\rangle}{E_m^{(0)} - E_k^{(0)}} \ .
\end{aligned} \tag{13.3.47}$$

The energy, to second order, is now given by

$$E_{m,\mu}^{(2)} = \sum_{k\neq m}\sum_{\kappa=1}^{g_k} a_{m\mu,k\kappa}^{(1)} \frac{\left|\langle\psi_{m,\mu}|H'|k,\kappa\rangle^{(0)}\right|^2}{E_m^{(0)} - E_k^{(0)}} \ . \tag{13.3.48}$$

We next illustrate the techniques just developed with some simple examples.

13.4 Example: Spin Hamiltonian

We want to consider the following simple "Hamiltonian"

$$H = AS_z^2 + b(S_x^2 - S_y^2) \tag{13.4.49}$$

for a system with a total spin $S = 1$. This Hamiltonian arises if one considers an ion with spin 1 located in a crystal at a point where the effective potential has rhombic symmetry and one only considers this perturbation on the background of large kinetic and Coulomb energies. The constants A and b are determined by the ionic crystal properties. In general $|b| < |A|$ so that we can treat the term $b(S_x^2 - S_y^2)$ as a perturbation on the term AS_z^2.

The eigenfunctions of the unperturbed Hamiltonian AS_z^2 are

$$|1\rangle^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad |2\rangle^{(0)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad , \quad |3\rangle^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{13.4.50}$$

with the corresponding unperturbed energies

$$E_1^{(0)} = A\hbar^2 \quad , \quad E_2^{(0)} = 0 \quad , \quad E_3^{(0)} = A\hbar^2 \ . \tag{13.4.51}$$

Clearly $E_1^{(0)}$ and $E_3^{(0)}$ are degenerate. In fact on the basis set above the Hamiltonian becomes

$$H = \hbar^2 \begin{pmatrix} A & 0 & b \\ 0 & 0 & 0 \\ b & 0 & A \end{pmatrix} \ . \tag{13.4.52}$$

We first give an exact solution and then use Rayleigh-Schrödinger perturbation theory

13.4.1 Exact Solution

To diagonalize the Hamiltonian H we simply solve

$$H|n\rangle = E_n|n\rangle \ . \tag{13.4.53}$$

With

$$|n\rangle = a_{1,n}|1\rangle^{(0)} + a_{2,n}|2\rangle^{(0)} + a_{3,n}|3\rangle^{(0)} \tag{13.4.54}$$

this requires that

$$\det(H - E_n 1) = 0 \tag{13.4.55}$$

and gives us the characteristic equation

$$(A\hbar^2 - E_n)^2 E_n - (b\hbar^2)^2 E_n = 0 \ . \tag{13.4.56}$$

Hence one value of E_n is $E_2 = 0$ and E_1, E_3 are given by:

$$(A\hbar^2 - E_n)^2 - (b\hbar^2)^2 = 0 \quad n = 1, 3 \tag{13.4.57}$$

or

$$\begin{aligned} E_1 &= A\hbar^2 + b\hbar^2 \\ E_2 &= 0 \\ E_3 &= A\hbar^2 - b\hbar^2 \ . \end{aligned} \tag{13.4.58}$$

As is easily checked the corresponding normalized eigenkets are

$$|1\rangle = \frac{1}{\sqrt{2}}\left[|1\rangle^{(0)} + |3\rangle^{(0)}\right] \ , \ |2\rangle = |2\rangle^{(0)} \ , \ |3\rangle = \frac{1}{\sqrt{2}}\left[|1\rangle^{(0)} - |3\rangle^{(0)}\right] . \tag{13.4.59}$$

13.4.2 Rayleigh-Schrödinger Solution

Since $E_1^{(0)}$ and $E_3^{(0)}$ are degenerate we must use the degenerate perturbation theory approach. This requires that we diagonalize the perturbation part H' of the Hamiltonian in the degenerate subspace. On the basis set of unperturbed states this portion of the Hamiltonian is

$$H'_{\text{degenerate}} = \hbar^2 \begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix} \ . \tag{13.4.60}$$

The corresponding eigenvalues are

$$E_1^{(1)} = b\hbar^2 \quad , \quad E_1^{(1)} = -b\hbar^2 \ . \tag{13.4.61}$$

Also the corresponding eigenkets which must be linear combinations of $|1\rangle^{(0)}$ and $|3\rangle^{(0)}$ and are easily found to be

$$\frac{1}{\sqrt{2}}\left[|1\rangle^{(0)} \pm |3\rangle^{(0)}\right]$$

respectively. These are the kets $|\psi_1\rangle$ and $|\psi_3\rangle$ which we have to use to do further perturbation calculations. To first order we then get

$$\begin{aligned} &E_1^{(1)} = \langle\psi_1|H'|\psi_1\rangle = b\hbar^2 \\ &E_3 = \langle\psi_3|H'|\psi_3\rangle = -b\hbar^2 \end{aligned} \tag{13.4.62}$$

and to this order the eigenkets are $|\psi_1\rangle$, $|2\rangle^{(0)}$, and $|\psi_3\rangle$. These results coincide with the exact solution. It is easily checked that the second order perturbations vanish. We do this only for E_1. Thus,

$$E_1^{(2)} = \sum_{k\neq 1,3} \frac{\left|\langle\psi_1|H'|k\rangle^{(0)}\right|^2}{E_1^{(0)} - E_k^{(0)}} = \frac{\left|\langle\psi_1|H'|2\rangle^{(0)}\right|^2}{E_1^{(0)} - E_2^{(0)}} = 0 \tag{13.4.63}$$

as stated.

13.5 Problems

13.1 Find the shift in the energy of the $n = 2$ levels of an hydrogen atom, to first order due to a constant electric field (linear Stark effect). The potential is

$$V' = -e\mathbf{E}\cdot\mathbf{r} = -e\mathcal{E}z \ .$$

13.2 A particle is in a 2-dimensional box of sides a. If a perturbation

$$V' = \lambda xy$$

is applied find the change in the energy of the ground state and first excited state to first non-trivial order. Find also the wavefunctions correct to first order.

13.3 For the two-dimensional simple harmonic oscillator with

$$H_0 = \hbar\omega(a_1^\dagger a_1 + a_2^\dagger a_2)$$

calculate the effect, to second order, of the perturbation

$$H' = \lambda(a_1^\dagger a_1^\dagger a_1 a_1 + a_2^\dagger a_2^\dagger a_2 a_2)$$

on the second excited states and to first order on the third excited states. What are the effects on the ground state and first excited states?

13.4 Repeat problem 13.3 with

$$H' = \lambda[(a_1^\dagger a_2 + a_2^\dagger a_1)]$$

This problem can also be solved exactly by introducing operators

$$A_1 = \cos\theta a_1 + \sinh\theta a_2$$

$$A_2 = -\sin\theta a_1 + \cos\theta_2$$

and choosing θ appropriately. Do this and compare with the perturbation result.

13.5 For a particle of mass m moving in the potential

$$V = \frac{1}{2}k_1x^2 + \frac{1}{2}k_2y^2 + \lambda xy \quad |k_1 - k_2| < 2\lambda$$

Find to order λ^2 the energy of the first exited state. Compare your answer with the exact solution obtained in problem 12.8c.

13.6 A particle of mass m and a charge q is placed in a box of sides (a, a, b) where $b < a$. A weak electric field

$$\mathbf{E} = \mathcal{E}(y/a\,, x/a\,, 0)$$

is applied to this particle. Find the energy of the ground state and first excited states correct to order $|\mathcal{E}|$.

13.7 Consider the Hamiltonian

$$H_0 = \frac{p^2}{2m} \quad \text{on} \quad -a \leq x \leq a$$

with the domain for p given by

$$D_p = \{f(x) \in C^2 \mid f(-a) = -f(a)\} \quad .$$

Find the shift in energy of all the energy levels to first nontrivial order due to a perturbation

$$\lambda H' = \lambda x \quad .$$

Hint: The solutions for H_0 are given in section 6.8.

13.8 A system with moment of inertia I has an Hamiltonian

$$H_0 = \frac{L^2}{2I} \quad .$$

a) What are the energies of the lowest and first excited states?
b) A perturbation

$$H' = g\frac{eB}{mc}L_x$$

is applied. Find the splitting of the first excited states.

13.9 Find the energy correct to order λ^2 for the second excited state of the Hamiltonian

$$H = H_0 + \lambda H'$$

where

$$H_0 = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{1}{2}m\omega^2(x^2 + y^2)$$

and

$$H' = xp_y \quad .$$

Can this problem be solved exactly?

Bibliography

[13.1] The book by Wigner [12.2], cited in the previous chapter, provides a very detailed treatment of degenerate Rayleigh-Schrödinger perturbation theory and shows how group-theoretic methods may be used to simplify the computations.

Chapter 14

Further Approximation Methods

14.1 Introduction

The approximation methods for the state vectors and energy levels of a given system that we discussed in the two previous chapters work very well indeed when the perturbation part H' of the Hamiltonian is small. The techniques were based on truncating the power series expansions for the state vector and energy. In practice it is often simpler to estimate the energy levels by assuming some approximate wavefunction with a reasonable shape. In this case one has to optimize the fit by varying the parameters that determine things like the width of the approximate wavefunction. In a case like this, the computations are greatly facilitated by the use of an expression that is stationary since the result will then be fairly insensitive to the details of the wavefunction.

The use of a stationary expression comes even more to the fore when one picks a set $\{\psi_\alpha\}$ of wavefunctions which depend on a set of parameters. Here α is a generic label for such a set of parameters. Since the expression used is to be stationary in the vicinity of the exact value we clearly expect the approximation to optimize by choosing those values of the parameters that do, in fact, make the expression stationary. In general, however, it is not enough to make the expression stationary. What one needs, in fact, is an extremum principle. We now elaborate these statements somewhat.

Suppose for example that a quantity to be calculated, say Q, is some functional of the wavefunction

$$Q = F(\psi)\,. \tag{14.1.1}$$

An example of such an expression would be the formula for the energy E given by

$$E = \frac{(\psi, H\psi)}{(\psi,\psi)}\,. \tag{14.1.2}$$

Suppose now that the expression (14.1.1) for Q is stationary when ψ satisfies the correct Schrödinger equation i.e. when we have the exact ψ. We then take our family $\{\psi_\alpha\}$ of wavefunctions and compute the quantities

$$Q(\alpha) = F(\psi_\alpha)\,. \tag{14.1.3}$$

By varying the parameters α we now make $Q(\alpha)$ stationary. Thus, we compute $\delta F/\delta\psi$ and form the set of equations

$$\frac{\delta F}{\delta \psi} = \frac{\partial Q}{\partial \alpha} = 0\,. \tag{14.1.4}$$

This will yield one (or several) values α' for the parameters α such that Q is stationary. One would then expect these values of α to optimize the value for Q. However, unless one has an *extremum* principle rather than just a *stationary* principle this will not necessarily be the case.

For example, let $F(\psi)$ be maximized by the correct ψ. Then choosing a ψ which satisfies (14.1.4) could, in fact, give the worst value for Q if it minimizes F. This shows the necessity for an extremum principle if one wants to be sure of optimizing the approximation.

In this chapter we first develop and then demonstrate the use of an extremum principle for the ground state energy of a system. The technique can also be extended to higher energy levels and we shall indicate how this is done. In general, however, the method is most useful for the ground state energy. That the ground state energy should satisfy a minimum principle is not at all surprising since it is after all the minimum energy that a system can possess.

In the last half of this chapter we introduce and discuss the "geometrical optics" approximation for quantum mechanics. This is also known as the Wentzel-Kramers-Brillouin or WKB approximation (see problems 2.7, 2.8), although many more names could be associated with this approximation. This then concludes our treatment of time independent approximation methods.

14.2 Rayleigh-Ritz Method

We shall not be concerned with the formal structure of the variational formulas obtained, nor shall we be concerned with their possible deeper interpretations. To us, the purpose of variational formulas, will be simply a means for facilitating approximations. As a first step we derive a minimum principle for the ground state energy for a system.

Theorem 1

Let H be the Hamiltonian of a system, then the functional,

$$E(\psi) = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} \tag{14.2.5}$$

is minimized when $|\psi\rangle$ is the ground state wavefunction.

Proof

Let $\{|n\rangle\}$ be the complete set of eigenkets of H corresponding to the eigenvalues

E_n where E_0 is the ground state energy, E_1 is the energy of the first excited state and so on. Now consider any normalized state

$$|\psi\rangle = \sum_n a_n |n\rangle \ . \tag{14.2.6}$$

Then,

$$E(\psi) = \sum_n E_n |a_n|^2 \tag{14.2.7}$$

or

$$E(\psi) \geq \sum_n E_0 |a_n|^2 = E_0 \tag{14.2.8}$$

where the equal sign in the "greater than or equal to" holds only if all $a_n = 0$ for $n \neq 0$. In that case, however, $|\psi\rangle = |0\rangle$ is the eigenfunction corresponding to E_0 as required.

Theorem 2

Let $E(\psi)$ be defined as in (14.2.5). Then any $|\psi\rangle$, for which $E(\psi)$ is stationary, is an eigenfunction of H.

Proof

Rewrite (14.2.5) in the form

$$E\langle\psi|\psi\rangle = \langle\psi|H|\psi\rangle \ . \tag{14.2.9}$$

Varying both sides with respect to ψ yields

$$\delta E\langle\psi|\psi\rangle + E\langle\psi|\delta\psi\rangle + E\langle\delta\psi|\psi\rangle = \langle\psi|H|\delta\psi\rangle + \langle\delta\psi|H|\psi\rangle \tag{14.2.10}$$

and if E is stationary (i.e. $\delta E = 0$) this yields

$$\langle\psi|H - E|\delta\psi\rangle + \langle\delta\psi|H - E|\psi\rangle = 0 \ . \tag{14.2.11}$$

Now although the variations $|\delta\psi\rangle$ and $\langle\delta\psi|$ are not independent they may be treated as independent since (14.2.11) is valid for arbitrary variations. In that case replacing $|\delta\psi\rangle$ by $i|\delta\psi\rangle$ in (14.2.11) yields

$$i\langle\psi|H - E|\delta\psi\rangle - i\langle\delta\psi|H - E|\psi\rangle = 0 \ . \tag{14.2.12}$$

Multiplying (14.2.12) by i and adding the result to and subtracting it from (14.2.11) yields

$$\langle\psi|H - E|\delta\psi\rangle = 0 \quad \text{and} \quad \langle\delta\psi|H - E|\psi\rangle = 0 \ . \tag{14.2.13}$$

These two equations are equivalent to (14.2.11) if the variations $|\delta\psi\rangle$ and $\langle\delta\psi|$ are treated as independent and arbitrary. In that case, however, they imply

$$(|H - E)|\psi\rangle = 0 \tag{14.2.14}$$

$$\langle\psi|(H^\dagger - E^*) = 0 \ . \tag{14.2.15}$$

On the other hand, H is an observable and therefore self-adjoint so that E is also real. Thus, (14.2.15) reads

$$\langle\psi|(H-E)=0\,. \tag{14.2.16}$$

This is simply the hermitean adjoint of (14.2.14). Thus, as stated, if E is stationary $|\psi\rangle$ satisfies the Schrödinger equation for this energy.

This now provides a means for obtaining a minimum principle for the higher energy levels. The technique presupposes that the wavefunctions for the lower levels are known. Thus, for example, if we are interested in finding the energy of the first excited state and if $|0\rangle$ is the exact ground state wavefunction then it is a simple matter to show that

$$E_1(\psi)=\frac{(\langle\psi|-\langle\psi|0\rangle\langle 0|)\,H\,(|\psi\rangle-\langle 0|\psi\rangle|0\rangle)}{(\langle\psi|-\langle\psi|0\rangle\langle 0|)\,(|\psi\rangle-\langle 0|\psi\rangle|0\rangle)} \tag{14.2.17}$$

is minimized by the wavefunction for the first excited state.

Similarly if $\{|k\rangle\ ,\ k=0,1,2,\ldots\}$ are the wavefunctions corresponding to the first n levels and

$$P_{n-1}=\sum_{k=0}^{n-1}|k\rangle\langle k| \tag{14.2.18}$$

then,

$$E_n(\psi)=\frac{\langle\psi|(1-P_{n-1})H(1-P_{n-1})|\psi\rangle}{\langle\psi|(1-P_{n-1})|\psi\rangle} \tag{14.2.19}$$

is minimized by $|\psi\rangle=|n\rangle$, the wavefunction for the nth excited state. As we stated before, this is not generally a useful method for obtaining the energy of excited states.

To use the variational method in practice one picks a set of wavefunctions $|\psi(\alpha_1,\alpha_2,\ldots,\alpha_n)\rangle$ depending on a set of n parameters $(\alpha_1,\alpha_2,\ldots,\alpha_n)$ and computes

$$\langle H\rangle=\frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}\ . \tag{14.2.20}$$

One then minimizes $\langle H\rangle$ with respect to these parameters. In choosing these trial wavefunctions one must insure that they satisfy the correct boundary conditions. What then remains is to solve the n equations

$$\frac{\partial\langle H\rangle}{\partial\alpha_j}=0\quad j=1,2,\ldots,n \tag{14.2.21}$$

for the n parameters α_j. If several sets of solutions are obtained one takes that set which minimizes $\langle H\rangle$. We now give some illustrations.

14.3 Example: Simple Harmonic Oscillator

This first example is simply to illustrate the technique before we consider a more realistic problem. We have

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \,. \tag{14.3.22}$$

As a trial wavefunction we choose

$$\psi(x) = \left(\frac{2\alpha}{\pi}\right)^{1/4} e^{-\alpha x^2} \,. \tag{14.3.23}$$

Even if we did not know the answer this choice would be reasonable since the ground state must be an even function of x (an even parity state) and has to decay rapidly at ±∞. Evaluating $\langle H\rangle$ we get

$$\langle H\rangle = \frac{\hbar^2\alpha}{2m} + \frac{m\omega^2}{8\alpha} \,. \tag{14.3.24}$$

Solving for the value of α that minimizes $\langle H\rangle$, namely the equation

$$\frac{\partial\langle H\rangle}{\partial\alpha} = 0 \tag{14.3.25}$$

we get

$$\alpha = \frac{m\omega}{2\hbar} \tag{14.3.26}$$

and hence the ground state energy is approximated by

$$\langle H\rangle|_{\alpha=(m\omega)/(2\hbar)} = \frac{1}{2}\hbar\omega \,. \tag{14.3.27}$$

Since we chose a wavefunction of the right form, this result coincides with the exact result as previously found. We next consider a more realistic example.

14.4 Example: He Ground State

We discussed this problem before when we did perturbation theory. The Hamiltonian is given by

$$H = \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} - 2e^2\left(\frac{1}{\mathbf{r}_1} + \frac{1}{\mathbf{r}_2}\right) + \frac{e^2}{\mathbf{r}_{12}} \,. \tag{14.4.28}$$

As we saw previously, if we neglect the term $e^2/\mathbf{r}_{12}$ arising from the interaction of the two electrons, we obtain an approximate wavefunction

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{8}{\pi a_0^3} e^{-2(r_1+r_2)/a_0} \,. \tag{14.4.29}$$

In fact, we used this wavefunction as a basis for our perturbation calculation. Now the effect of the interaction of the two electrons is to repel each other. This can be thought of as a mutual screening of the nucleus. In that case it becomes reasonable to think of the range of the wavefunction as a parameter. Thus, we take as a normalized trial wavefunction

$$\psi(\mathbf{r}_1,\mathbf{r}_2) = \frac{Z^3}{\pi a_0^3} e^{-Z(r_1+r_2)/a_0} \tag{14.4.30}$$

where Z is our variational parameter. With this wavefunction we evaluate

$$\langle H\rangle = \langle\psi|H|\psi\rangle \tag{14.4.31}$$

just as we did in the perturbation calculation to get (see problem 14.2)

$$\langle H\rangle = \frac{e^2}{a_0}\left(Z^2 - \frac{27}{8}Z\right) \ . \tag{14.4.32}$$

Solving $\partial\langle H\rangle/\partial Z = 0$ yields:

$$Z = \frac{27}{16} \ . \tag{14.4.33}$$

Thus, our physical intuition was correct and Z is indeed less than 2 and greater than 1, indicating a screened nucleus. Substituting this result back in (14.4.32) we get an estimate for the ground state energy

$$E_0 = -\frac{e^2}{a_0}\left(\frac{27}{16}\right)^2 = -76.6\ \mathrm{eV} \ . \tag{14.4.34}$$

This is closer to the experimental values of -78.6 eV than the value of -74.0 eV previously obtained from perturbation theory. Furthermore both approximate values are higher than the experimental result. This is simply in conformity with the minimum principle previously enunciated.

The Rayleigh-Ritz variational principle elucidated above has found a tremendous amount of application in molecular physics. The use of computers has facilitated the use of trial functions with very large numbers of parameters. Although one can thus improve the fits to experimental data, the physical interpretation of many of the parameters is frequently obscure. This situation is particularly acute in these cases since one is not only trying to obtain the best fit for the energy eigenvalue but is actually interested in the wavefunction itself. Nevertheless such unesthetic procedures are sometimes unavoidable.

14.5 The WKB Approximation

In Chapter 2, part 4 we gave a heuristic derivation of the Schrödinger equation by postulating that the Hamilton-Jacobi equation is a geometrical optics approximation for some wave equation. This was examined a little further in problems 2.7 and 2.8. We now study this question more systematically to develop a useful approximation technique.

It is useful to recall the geometrical optics approximation and work at first in analogy to it. We then justify our procedure later. Now, in general, geometrical optics is valid if the index of refraction n varies little in a distance of one wavelength. This means we need

$$\lambda\left|\frac{dn}{dx}\right| < 1\,. \tag{14.5.35}$$

Actually a number of step discontinuities in n are permitted as long as the distances between these discontinuities are also large compared to the wavelength. This occurs for instance at the surfaces of mirrors or lenses in optics. On the other hand, if the jump discontinuities are separated by distances comparable to the wavelength, as in the case of gratings, physical optics becomes necessary.

In Chapter 1 we saw that if an index of refraction is to be associated with a given medium for matter waves then $n \propto p$ and since $\lambda \propto 1/p$ condition (14.5.35) reduces to

$$\left|\lambda\frac{d}{dx}\left(\frac{\lambda_0}{\lambda}\right)\right| = \frac{\lambda_0}{\lambda}\left|\frac{d\lambda}{dx}\right| < 1 \tag{14.5.36}$$

where we have written $n = \lambda_0/\lambda$. In general n is of the order of unity so this condition reduces to

$$\left|\frac{d\lambda}{dx}\right| < 1\,. \tag{14.5.37}$$

Thus, we expect a geometrical optics approximation for quantum mechanics to be valid when (14.5.37) is satisfied. Now,

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2m(E-V)}}\;. \tag{14.5.38}$$

Therefore,

$$\left|\frac{d\lambda}{dx}\right| = \frac{mh\left|\frac{dV}{dx}\right|}{|2m(E-V)|^{3/2}}\;. \tag{14.5.39}$$

So (14.5.37) becomes

$$\frac{mh\left|\frac{dV}{dx}\right|}{|2m(E-V)|^{3/2}} < 1\;. \tag{14.5.40}$$

Having given a heuristic argument for the validity of a geometrical optics approximation we now derive the approximation as well as its region of validity.

In the transition from physical to geometrical optics we ignore the wave nature of light and follow the trajectories of light rays. Thus, to make the analogous transition for particles is to ignore their wave nature and follow their trajectories. We accomplish this by expanding in powers of $\hbar$ and keeping only the lowest order terms.

Our starting point is the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi\,. \tag{14.5.41}$$

First we write

$$\Psi = A_0 e^{iw/\hbar} \tag{14.5.42}$$

with A_0 a constant. The Schrödinger equation (14.5.41) then becomes

$$\frac{\partial w}{\partial t} + \frac{1}{2m}(\nabla w)^2 + V - \frac{i\hbar}{2m}\nabla^2 w = 0\ . \tag{14.5.43}$$

Letting $\hbar \to 0$ we get

$$\frac{\partial w}{\partial t} + H(\mathbf{r}, \nabla w) = 0\ . \tag{14.5.44}$$

This is just one form of the classical Hamilton-Jacobi equations (equation (2.4.61)) where w is Hamilton's principal function. Furthermore, just as in classical mechanics one can get a separation of the time variable. Thus, if Ψ is a stationary state, we have

$$\Psi = \psi(\mathbf{r}) e^{-iEt/\hbar}\ . \tag{14.5.45}$$

In that case

$$w(\mathbf{r}, t) = S(\mathbf{r}) - Et \tag{14.5.46}$$

where $S(\mathbf{r})$ corresponds to Hamilton's characteristic function. We now have

$$\psi(\mathbf{r}) = A_0 e^{-iS(\mathbf{r})/\hbar} \tag{14.5.47}$$

and

$$\frac{1}{2m}(\nabla S)^2 - (E - V) - \frac{i\hbar}{2m}\nabla^2 S = 0\ . \tag{14.5.48}$$

The WKB approximation uses the last term in (14.5.48) to get one term beyond the classical expression.

We now show how to do this quite generally for the one-dimensional Schrödinger equation. In one dimension we start with

$$\frac{d^2u}{dx^2} + \frac{2m}{\hbar^2}(E - V)u = 0 \tag{14.5.49}$$

and put as before,

$$u = Ae^{-iS/\hbar}\ . \tag{14.5.50}$$

However, in order that A and S be real we also let A depend on x. Substituting (14.5.50) into (14.5.49) we get on separating real and imaginary parts

$$\left(\frac{dS}{dx}\right)^2 - 2m(E - V) = \hbar^2 \frac{1}{A}\frac{d^2A}{dx^2} \tag{14.5.51}$$

and

$$2\frac{dA}{dx}\frac{dS}{dx} + A\frac{d^2S}{dx^2} = 0\ . \tag{14.5.52}$$

From (14.5.52) we get, using primes to indicate differentiation, that

$$\frac{A'}{A} = -\frac{1}{2}\frac{S''}{S'} \tag{14.5.53}$$

so that

$$A = A_0 (S')^{-1/2} . \tag{14.5.54}$$

Substituting this back in (14.5.51) we find:

$$(S')^2 = 2m(E - V) + \hbar^2 \left[\frac{3}{4} \left(\frac{S''}{S'} \right)^2 - \frac{S'''}{S'} \right] . \tag{14.5.55}$$

Now writing S as a power-series expansion in $\hbar^2$

$$S = S_0 + \hbar^2 S_1 + \hbar^4 S_2 + \dots \tag{14.5.56}$$

and collecting the coefficients of equal powers of $\hbar^2$ we get to 0th order in $\hbar^2$

$$(S_0')^2 = 2m(E - V) \tag{14.5.57}$$

and to first order in $\hbar^2$

$$2S_0' S_1' = \frac{3}{4} \left(\frac{S_0''}{S_0'} \right)^2 - \frac{1}{2}\frac{S_0''}{S_0'} . \tag{14.5.58}$$

The S_1 term gives a correction to the S_0 term. We use it to estimate the region of validity of the WKB approximation, but first we solve (14.5.57) for S_0. To integrate this equation we must distinguish between two regions.

1. The *classically allowed* region $E > V$.

2. The *classically forbidden* region $E < V$.

1. In the classically allowed region we define

$$k(x) = \frac{1}{\hbar}\sqrt{2m(E - V(x))} . \tag{14.5.59}$$

Then

$$S_0' = \pm \hbar k(x) \tag{14.5.60}$$

giving

$$S_0 = \pm \hbar \int_{x_1}^{x} k(x')\, dx' + \varphi \tag{14.5.61}$$

and combining this with (14.5.54) we get the approximate wave function

$$u_0 = \frac{A_0}{\sqrt{k(x)}} \cos \left[\int_{x_1}^{x} k(x')\, dx' + \varphi \right] \tag{14.5.62}$$

where A_0 and φ are constants and the lower limit x_1 of the integral may be picked to be any convenient number.
2. In the classically forbidden region we define

$$\kappa(x) = \frac{1}{\hbar}\sqrt{2m(V(x) - E)}\,. \tag{14.5.63}$$

leading eventually to the approximate wave function

$$u_0 = \frac{1}{\sqrt{\kappa(x)}}\left[B_0 \exp\int_{x_1}^{x} \kappa(x')\,dx' + C_0 \exp - \int_{x_1}^{x} \kappa(x')\,dx'\right]\ . \tag{14.5.64}$$

Before proceeding with the discussion of how to match the solutions (14.5.62), (14.5.64) at the classical turning points given by $E = V(x)$ we examine conditions for the validity of this approximation. In fact, we only derive a necessary condition, but in practice this condition is also usually sufficient.

Recall that by analogy with optics we arrived at the condition for validity of geometrical optics that

$$\frac{mh\left|\frac{dV}{dx}\right|}{|2m(E-V)|^{3/2}} < 1\,. \tag{14.5.65}$$

Now in order that u_0 be a good approximation it is necessary that the next correction to u_0 be small. The next correction gives a factor $\exp i\hbar S_1$ and this will be close to unity, so that the correction will be small, if $\hbar S_1$ is small. Thus, a necessary condition is that

$$|\hbar S_1| < 1\,. \tag{14.5.66}$$

We therefore look at (14.5.58), the equation for S_1 in more detail. Now in the classically allowed region $(E > V)$, $S_0' = \pm\hbar k(x)$ so that substituting this result into (14.5.58) yields

$$\begin{aligned} S_1' &= \pm\frac{1}{4\hbar}\left[\frac{3k'^2}{2k^3} - \frac{k''}{k^2}\right] \\ &= \mp\frac{1}{4\hbar}\left[\frac{3}{2}\left(\frac{k'}{k^2}\right)' + \frac{1}{2}k\left(\frac{k'}{k^2}\right)^2\right]\,. \end{aligned} \tag{14.5.67}$$

Hence,

$$\hbar S_1 = \mp\frac{1}{4}\left[\frac{k'}{k^2} + \frac{1}{2}\int_{x_1}^{x}\left(\frac{k'}{k^2}\right)^2 k(x')\,dx'\right]\ . \tag{14.5.68}$$

Clearly if $|k'/k^2| < 1$ then $\hbar S_1 < 1$ as required. This, however, is precisely the condition (14.5.40) or (14.5.65). Thus, our intuitive formula is verified. For the classically forbidden region $(E < V)$ simply replace k by $i\kappa$ to get a similar result. Both results can be included in the single condition

$$\left|\frac{m\hbar\frac{dV}{dx}}{2m(E-V)^{3/2}}\right| < 1\,. \tag{14.5.69}$$

Although the condition (14.5.69) is necessary in order that the first-order WKB approximation be useful, it tells us nothing about the convergence of the power series for S. Detailed examination shows that this series is asymptotic and therefore provides a good approximation if $\hbar$ is small. This simply means that, in examining the series, $\hbar$ is treated as a parameter and not that $\hbar$ actually varies. In practice, if (14.5.69) is satisfied, $\hbar$ is sufficiently small so that the first term yields a good approximation.

14.5.1 Turning points

Those points x at which

$$E = V(x) \tag{14.5.70}$$

are where the kinetic energy changes sign. Classically of course, the kinetic energy is always non-negative and therefore the classical motion reverses at these points. These points are known, therefore, as turning points. Furthermore, the approximate wave function u_0 given by (14.5.62) and (14.5.64) changes at a turning point. This "discontinuity" is not a property of the solutions of the original Schrödinger equation (14.5.49) but is rather a consequence of the approximations made. In fact, at the turning point we have that $k = 0$, so that the condition $|k'/k^2| < 1$, necessary for the validity of the approximation, is violated. Thus, the approximate solutions u_0 are valid only up to some distance (several wavelengths) from the turning points. It is therefore necessary to find a means of connecting an approximate solution in the classically forbidden region with an approximate solution in the classically allowed region. The difficulty just mentioned can be traced back to the effective potential for the approximate solutions. By straightforward differentiation we find that in the classically allowed region u_0 satisfies the following Schrödinger equation

$$\frac{d^2u_0}{dx^2} + \left(k^2 - \frac{3k'^2}{4k^2} + \frac{k''}{2k}\right) u_0 = 0 \ . \tag{14.5.71}$$

Thus, we have in fact introduced a singular effective potential. The singularities, being due to the vanishing of k, occur precisely at the turning points.

Now suppose that x_1 is a turning point with the allowed region $x > x_1$ (see fig. 14.1). In the region of the turning point we can approximate the potential $V(x)$ by the tangent to $V(x)$ at $x = x_1$. Thus, we have

$$V(x) \approx E - \frac{\hbar^2}{2m}c^2(x - x_1) \quad \text{near} \quad x = x_1 \ . \tag{14.5.72}$$

This leads to

$$\begin{aligned}
k^2(x) &= \frac{2m}{\hbar^2}(E - V) = c^2(x - x_1) \qquad x > x_1 \\
\kappa^2(x) &= \frac{2m}{\hbar^2}(V - E) = -c^2(x - x_1) \quad x < x_1 \ .
\end{aligned} \tag{14.5.73}$$

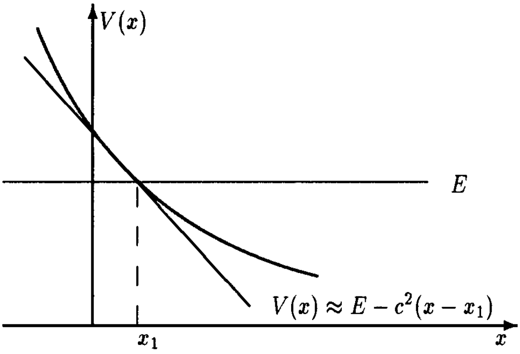

Figure 14.1: Turning point and linear approximation for potential.

Substituting these linear approximations for k^2 and κ^2 into the Schrödinger equation we obtain

$$\begin{aligned} \frac{d^2u}{dx^2} + c^2(x - x_1)u &= 0 \quad x > x_1 \\ \frac{d^2u}{dx^2} + c^2(x_1 - x)u &= 0 \quad x < x_1 \ . \end{aligned} \tag{14.5.74}$$

The solutions of these equation can be expressed in terms of Bessel functions of order 1/3.

$$\begin{aligned} u_a^{\pm}(x) &= A_{\pm} y^{1/3} J_{\pm 1/3}(y) \quad x > x_1 \\ u_f^{\pm}(x) &= B_{\pm} z^{1/3} I_{\pm 1/3}(z) \quad x < x_1 \end{aligned} \tag{14.5.75}$$

where "a" stands for "allowed" ($x > x_1$ in this case) and "f" stands for "forbidden" ($x < x_1$ in this case), and

$$\begin{aligned} y &= \int_{x_1}^{x} k(x')\, dx' = \frac{2}{3} c(x - x_1)^{3/2} \quad x > x_1 \\ z &= \int_{x}^{x_1} \kappa(x')\, dx' = \frac{2}{3} c(x_1 - x)^{3/2} \quad x < x_1 \ . \end{aligned} \tag{14.5.76}$$

As stated, $J_n(x)$ is an ordinary Bessel function satisfying equation (10.3.33) and $I_n(x)$ is a modified Bessel function corresponding to a $J_n(x)$ with imaginary argument.

These, equations (14.5.75), are exact solutions of a differential equation that coincides with the exact Schrödinger equation at a turning point whenever (14.5.72) is a good approximation. Thus, we can use these solutions to match the WKB solutions across a turning point. Away from a turning point we have the WKB solutions. Also away from the turning point we can use the asymptotic forms of the solutions (14.5.75). Notice that both $y(x)$ and

$z(x)$ were defined so as to increase as we move away from the turning point. The asymptotic forms of these solutions can be found in any book on Bessel or transcendental functions [14.2] and are given by

$$J_{\pm 1/3}(y) \to \sqrt{\frac{2}{\pi y}}\cos(y - \pi/4 \mp \pi/6) \quad \text{for} \quad y \to \infty \tag{14.5.77}$$

$$I_{\pm 1/3}(z) \to \sqrt{\frac{1}{2\pi z}}\left[e^{z} + e^{-z}e^{-i\pi(1/2\pm 1/3)}\right] \quad \text{for} \quad z \to \infty\ . \tag{14.5.78}$$

The exponentially damped term e^{-z} in (14.5.78) is only meaningful if we take a linear combination $I_{1/3}(z) - I_{-1/3}(z)$ so that the e^{z} terms cancel, since in writing (14.5.78) we have dropped terms of order e^{z}/z and these are large compared to e^{-z}. For small $|x - x_1|$ we also have the asymptotic forms

$$J_{\pm 1/3}(y) \to \frac{(y/2)^{\pm 1/3}}{\Gamma(1 \pm 1/3)} \quad \text{for} \quad y \to 0 \tag{14.5.79}$$

$$I_{\pm 1/3}(z) \to \frac{(z/2)^{\pm 1/3}}{\Gamma(1 \pm 1/3)} \quad \text{for} \quad z \to 0\ . \tag{14.5.80}$$

Thus, in the vicinity of a turning point the solutions (14.5.75) become

$$\begin{aligned} u_a^+ &\approx A_+\frac{2^{-1/3}(2c/3)^{2/3}}{\Gamma(4/3)}(x - x_1) && (14.5.81)\\ u_a^- &\approx A_-\frac{2^{1/3}}{\Gamma(2/3)} && (14.5.82)\\ u_f^+ &\approx B_+\frac{2^{-1/3}(2c/3)^{2/3}}{\Gamma(4/3)}|x - x_1| && (14.5.83)\\ u_f^- &\approx B_-\frac{2^{1/3}}{\Gamma(2/3)}\ . && (14.5.84) \end{aligned}$$

So we find that u_a^+ joins smoothly to u_f^+ if $A_+ = -B_+$, and u_a^- joins smoothly to u_f^- if $A_- = B_-$. So we choose the constants

$$A_+ = -B_+ = A_- = B_- \equiv A\ . \tag{14.5.85}$$

The solutions then divide into two kinds: u^+ and u^-. They join smoothly and have the asymptotic forms
<u>allowed region</u>

$$u_a^+ \to A\sqrt{\frac{2}{\pi k(x)}}\cos(y - 5\pi/12) \quad \text{for} \quad |x - x_1| \to \infty \tag{14.5.86}$$

<u>forbidden region</u>

$$u_f^+ \to -A\frac{1}{\sqrt{2\pi\kappa(x)}}\left[e^{z} + e^{-z}e^{-i\pi 5/6}\right] \quad \text{for} \quad |x - x_1| \to \infty \tag{14.5.87}$$

and
allowed region

$$u_a^- \to A\sqrt{\frac{2}{\pi k(x)}}\cos(y-\pi/12) \quad \text{for} \quad |x-x_1| \to \infty \tag{14.5.88}$$

forbidden region

$$u_f^- \to A\frac{1}{\sqrt{2\pi\kappa(x)}}\left[e^z + e^{-z}e^{-i\pi/6}\right] \quad \text{for} \quad |x-x_1| \to \infty \ . \tag{14.5.89}$$

Thus, for example, if we want an asymptotically damped solution in the classically forbidden region we take u_f^+ plus u_f^-. In the classically forbidden region this has the asymptotic form

$$A\frac{1}{\sqrt{2\pi\kappa}}\left[e^{-i\pi/6} - e^{-i\pi 5/6}\right]e^{-z} \tag{14.5.90}$$

and joins smoothly onto the solution $u_a^+ + u_a^-$ which in the classically allowed region has the asymptotic behaviour

$$A\sqrt{\frac{2}{\pi k}}\left[\cos(y-5\pi/12)+\cos(y-\pi/12)\right] = A\sqrt{\frac{2}{\pi k}}\sqrt{3}\cos(y-\pi/4) \ . \tag{14.5.91}$$

Simplifying this expression, by cancelling common factors, we obtain the first of our connection formulas

$$\kappa^{-1/2}e^{-z} \Rightarrow 2k^{-1/2}\cos(y-\pi/4) \ . \tag{14.5.92}$$

The fancy arrow means that the asymptotic solution in the classically forbidden region appearing on the left of the arrow goes over into the solution with asymptotic behaviour appearing on the right of the arrow. The converse is *false* since a very small error in the phase of the cosine term would integrate back to produce a small admixture of the dominating exponentially growing term on the left. In fact in computer integration of the Schrödinger equation one frequently integrates "in from infinity" precisely because one wants to avoid the exponentially growing term due to round-off error.

On the other hand if we want a solution that in the classically allowed region has a definite phase, as for example $k^{-1/2}\cos(y-\pi/4+\varphi)$ where $\varphi \neq n\pi$ with $n = 0, \pm 1, \pm 2, \pm \ldots$ we start by again realizing that

$$\begin{aligned}
u_a^+ + u_a^- &\to A\sqrt{\frac{2}{\pi k}}\left[\cos(y-5\pi/12)+\cos(y-\pi/12)\right] \\
&= A\sqrt{\frac{2}{\pi k}}2\cos(y-\pi/4)\cos(\pi/6) \\
&= A\sqrt{\frac{2}{\pi k}}\sqrt{3}\cos(y-\pi/4) \\
u_a^+ - u_a^- &\to A\sqrt{\frac{2}{\pi k}}\left[\cos(y-5\pi/12)-\cos(y-\pi/12)\right]
\end{aligned}$$

$$\begin{aligned} &= A\sqrt{\frac{2}{\pi k}}2\sin(y-\pi/4)\sin(\pi/6) \\ &= A\sqrt{\frac{2}{\pi k}}\sin(y-\pi/4) \end{aligned} \qquad (14.5.93)$$

and we further recall that

$$\cos(y-\pi/4)\cos\varphi - \sin(y-\pi/4)\sin\varphi = \cos(y-\pi/4+\varphi) \ . \qquad (14.5.94)$$

Thus, the solution in the classically forbidden region matching onto the one with the above asymptotic behaviour is given by:

$$\frac{1}{\sqrt{3}}\cos\varphi[u^+ + u^-] - \sin\varphi[u^+ - u^-] \qquad (14.5.95)$$

and assumes after some simplification the form

$$\sqrt{\frac{2}{\pi\kappa}}\sin\varphi\, e^z + \frac{1}{\sqrt{2\pi\kappa}}e^{-z+i\varphi} \ . \qquad (14.5.96)$$

The exponentially damped term on the right is clearly not meaningful and is to be dropped. Thus, we get another connection formula:

$$\frac{1}{\sqrt{k}}\cos(y-\pi/4+\varphi) \Rightarrow \frac{1}{\sqrt{\kappa}}\sin\varphi\, e^z \qquad \varphi \neq 0 \ . \qquad (14.5.97)$$

The two formulas, (14.5.92), (14.5.97), suffice for handling all problems. In using these formulas it is important to remember two things, namely that y and z are so defined as to increase as we move away from the turning point and secondly that the formulas may only be used to connect the solution on the left of the arrow to the solution on the right of the arrow and never in the reverse direction.

Next we illustrate the use of this method with some examples that frequently occur in practice.

14.6 WKB Applied to a Potential Well

Consider a potential well as shown in figure 14.2 rising to infinity for very large $|x|$. For the energy E shown there are clearly two classical turning points indicated by x_1 and x_2 . Regions 1 and 3 are classically forbidden and region 2 is classically allowed. Therefore, we want exponentially damped solutions in regions 1 and 3. As usual this should, and will, lead to an energy quantization.

Considering first the point x_1 we have using (14.5.62), (14.5.64) that

$$\begin{aligned} u_0 &= \frac{A}{\sqrt{\kappa(x)}}\exp - \int_x^{x_1}\kappa(x')\,dx' \quad x < x_1 \\ u_0 &= \frac{2A}{\sqrt{k(x)}}\cos\left[\int_{x_1}^{x} k(x')\,dx' - \frac{\pi}{4}\right] \quad x_1 < x < x_2 \ . \end{aligned} \qquad (14.6.98)$$

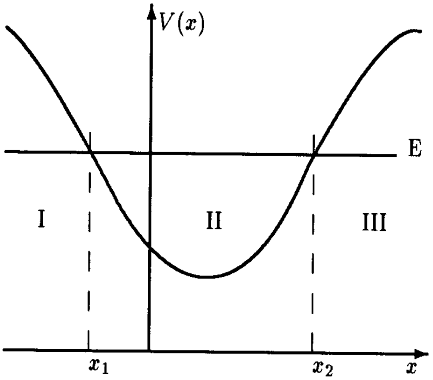

Figure 14.2: Bound state problem in WKB approximation.

Here we have already fixed the phase and amplitude, in the solution (14.6.98) for the classically allowed region 2, by applying the connection formula (14.5.92) at $x = x_1$. Similarly we have that in region 3

$$u_0 = \frac{B}{\sqrt{\kappa(x)}} \exp - \int_{x_2}^{x} \kappa(x')\,dx' \quad x > x_2 \tag{14.6.99}$$

and hence by applying the connection formula (14.5.92) at x_2 we get that for $x_1 < x < x_2$

$$\begin{aligned} u_0 &= \frac{2B}{\sqrt{k(x)}} \cos\left[\int_{x}^{x_2} k(x')\,dx' - \frac{\pi}{4}\right] \\ &= \frac{2B}{\sqrt{k(x)}} \cos\left[\int_{x}^{x_1} k(x')\,dx' - \frac{\pi}{4} + \int_{x_1}^{x_2} k(x')\,dx'\right] . \end{aligned} \tag{14.6.100}$$

This last solution can be further rewritten as

$$\begin{aligned} u_0 &= \frac{2B}{\sqrt{k(x)}} \cos\left[-\int_{x_1}^{x} k(x')\,dx' + \frac{\pi}{4} + \int_{x_1}^{x_2} k(x')\,dx' - \frac{\pi}{2}\right] \\ &= \frac{2B}{\sqrt{k(x)}} \cos\left[\int_{x_1}^{x} k(x')\,dx' - \frac{\pi}{4} - \alpha\right] \quad x_1 < x < x_2 \end{aligned} \tag{14.6.101}$$

where

$$\alpha = \int_{x_1}^{x_2} k(x')\,dx' - \frac{\pi}{2} . \tag{14.6.102}$$

Since this solution must coincide with the solution (14.6.98) for this region we require that

$$\alpha = n\pi \quad n = 0, 1, 2, \ldots \tag{14.6.103}$$

and

$$B = (-1)^n A \,. \tag{14.6.104}$$

Hence, we get the energy quantization condition

$$\int_{x_1}^{x_2} k(x')\,dx' = (n + 1/2)\pi \quad n = 0, 1, 2, \ldots \,. \tag{14.6.105}$$

But,

$$k(x) = \frac{1}{\hbar}\sqrt{2m(E - V)} = \frac{2\pi}{h}p \tag{14.6.106}$$

so that (14.6.103) reads

$$\int_{x_1}^{x_2} p\,dx = (n + 1/2)\frac{h}{2} \tag{14.6.107}$$

or more compactly

$$\oint p\,dx = (n + 1/2)h \,. \tag{14.6.108}$$

Except for the additional term $h/2$ on the right this is just the Bohr-Sommerfeld quantization rule which now emerges quite naturally as a semi-classical approximation to the Schrödinger equation.

14.6.1 Special Boundaries

So far we have only discussed turning points occurring at points where $V(x)$ is continuous. If $V(x)$ experiences a finite jump discontinuity then one knows that the exact solution of the Schrödinger equation as well as the first derivative are continuous. In this case one simply matches the function and its derivative at the point of discontinuity regardless of whether or not this point is also a turning point. As long as $V(x)$ varies slowly on either side of this point, the asymptotic WKB solutions may be used right up to it. In general this simplifies the calculation since it eliminates the necessity for explicit use of the connection formulas.

If $V(x)$ experiences an infinite discontinuity at the point $x = a$ so that say $V = \infty$ for $x < a$, then the exact wavefunction vanishes there and the appropriate WKB solution to use for $x > a$, a classically allowed region, is

$$u = \frac{1}{\sqrt{k(x)}}\sin\left(\int_a^x k(x')\,dx'\right) \,. \tag{14.6.109}$$

This assumes of course that $k(x)$ is slowly varying so that the whole WKB procedure is valid.

This type of situation occurs, for example, in practice for the s-wave ($l = 0$) radial wavefunction $R_0(r)$. One has an equation for $u(r) = rR_0(r)$ so that $u(r)$ must vanish at the origin.

14.7 WKB Approximation for Tunneling

As a second example to illustrate the use of the connection formulas we consider tunneling through a smooth potential barrier as shown in fig. 14.3. The turning points are x_1 and x_2 with $x_1 < x_2$. The classically allowed regions are
region I : $x < x_1$
and
region III : $x > x_2$.
The classically forbidden region is
region II : $x_1 < x < x_2$.

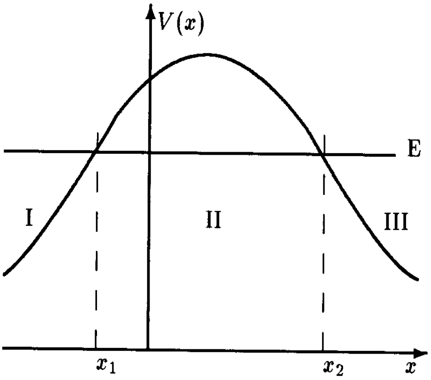

Figure 14.3: Tunneling in WKB approximation.

We are interested in the case of a particle with energy E incident from the left. Thus, in region I we have the WKB solution

$$\begin{aligned}\psi_I(x) &= \frac{1}{\sqrt{k(x)}}\left[\exp -i\left(\int_x^{x_1} k(x')\,dx' - \frac{\pi}{4}\right)\right.\\ &+ \left. R\exp i\left(\int_x^{x_1} k(x')\,dx' - \frac{\pi}{4}\right)\right] \,. \qquad (14.7.110)\end{aligned}$$

We have arbitrarily chosen the amplitude of the incoming wave as 1 and called the amplitude of the reflected wave R. Also the phase factors of $\exp(\pm i\pi/4)$ have been included to facilitate the use of the connection formulas. Rewriting (14.7.110) in terms of trigonometric functions we obtain

$$\begin{aligned}\psi_I(x) &= \frac{1}{\sqrt{k(x)}}\left[(1+R)\cos\left(\int_x^{x_1} k(x')\,dx' - \frac{\pi}{4}\right)\right.\\ &- \left. i(1-R)\sin\left(\int_x^{x_1} k(x')\,dx' - \frac{\pi}{4}\right)\right] \,. \qquad (14.7.111)\end{aligned}$$

In region II the general solution is given by:

$$\begin{aligned}\psi_{II}(x) &= \frac{1}{\sqrt{\kappa(x)}}\left[A\exp-\left(\int_{x_1}^{x}\kappa(x')\,dx'\right)\right.\\ &+ \left.B\exp\left(\int_{x_1}^{x}\kappa(x')\,dx'\right)\right] \ . \qquad (14.7.112)\end{aligned}$$

In region III we must have a purely transmitted wave. Hence, we get

$$\psi_{III}(x) = T\frac{1}{\sqrt{k(x)}}\exp i\left(\int_{x_2}^{x}k(x')\,dx' - \frac{\pi}{4}\right) \qquad (14.7.113)$$

where we have called the amplitude of this wave T. This wave can be rewritten as

$$\begin{aligned}\psi_{III}(x) &= T\frac{1}{\sqrt{k(x)}}\left[\cos\left(\int_{x_2}^{x}k(x')\,dx' - \frac{\pi}{4}\right)\right.\\ &+ \left. i\sin\left(\int_{x_2}^{x}k(x')\,dx' - \frac{\pi}{4}\right)\right] \ . \qquad (14.7.114)\end{aligned}$$

We now use the connection formula (14.5.92) at x_1 with

$$y = \int_{x}^{x_1}k(x')\,dx' \quad , \quad z = \int_{x_1}^{x}\kappa(x')\,dx' \qquad (14.7.115)$$

as well as (14.5.97) with $\varphi = -\pi/4$. This yields that

$$A = \frac{1}{2}(1+R) \quad , \quad B = i(1-R) \ . \qquad (14.7.116)$$

We repeat this procedure at x_2, but this time we have:

$$y = \int_{x_2}^{x}k(x')\,dx' \quad , \quad z = \int_{x}^{x_2}\kappa(x')\,dx' \ . \qquad (14.7.117)$$

So, we define

$$S = \exp-\int_{x_1}^{x_2}\kappa(x')\,dx' \ . \qquad (14.7.118)$$

Then ψ_{II} can be written to read:

$$\begin{aligned}\psi_{II}(x) &= \frac{1}{\sqrt{\kappa(x)}}\left[AS\exp\left(\int_{x}^{x_2}\kappa(x')\,dx'\right)\right.\\ &+ \left.BS^{-1}\exp\left(-\int_{x}^{x_2}\kappa(x')\,dx'\right)\right] \ . \qquad (14.7.119)\end{aligned}$$

The connection formulas can now be applied directly and yield

$$T = 2BS^{-1} = 2i(1-R)S^{-1} \qquad (14.7.120)$$

and

$$T = iAS = \frac{i}{2}(1 + R)S\,. \tag{14.7.121}$$

Solving these equations for T and R we find

$$T = \frac{iS}{1 + S^2/4} \approx iS \tag{14.7.122}$$

$$R = \frac{1 - S^2/4}{1 + S^2/4} \approx 1 - S^2/2\,. \tag{14.7.123}$$

The last two approximations are in keeping with the spirit of the WKB approximation which requires that $|S| < 1$. The quantities T and R have the obvious interpretation of transmission and reflection amplitude respectively.

Two more comments are in order regarding this example. The first deals with the region of validity of our approach. Clearly if $x_2 - x_1 = L$ is too small then the wavefunction ψ_{II} never assumes its asymptotic form and our connection formulas are not valid. This occurs if the energy E is too close to a maximum of the tunneling potential.

The second point is that we do not have connection formulas connecting an exponentially growing solution in the forbidden region to solutions in the allowed region. The reason is that the phase in the allowed region depends crucially on the admixture of the negligible, damped solution. Thus, if the forbidden region is of finite extent so that both exponentially growing and damped solutions can exist, then we must use both connection formulas. The exponentially damped solution in the forbidden region connects to the cosine solution with the phase $-\pi/4$ and amplitude 2 in the allowed region, whereas the other solution of definite phase in the classically allowed region connects to the exponentially growing solution in the classically forbidden region.

14.8 Alpha Decay

An interesting application of the WKB approximation is to alpha decay. For this process the experimental data show that half-lives of the alpha emitters vary from 10^{-7} seconds up to 10^{10} years while the energy of the emitted alphas varies only over a very small range from about 8.95 MeV for Po^{212} to 4.05 MeV for Th^{232}. The simple model we present is able to account for these extremes in half-lives.

We first present a somewhat heuristic discussion and then follow with a more detailed analysis based on a discussion by Gamow [14.4].

14.8.1 Heuristic Discussion

We begin by approximating the potential experienced by the alpha particle as shown in figure 14.4. We consider only the s-wave and write the radial

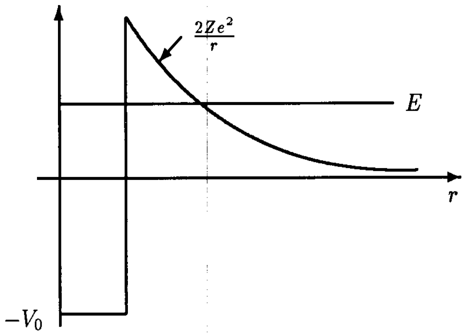

Figure 14.4: Model potential for alpha decay.

wavefunction in the form

$$\psi(r) = \frac{u(r)}{r} . \tag{14.8.124}$$

Then, $u(r)$ satisfies

$$\frac{d^2u}{dr^2} = \begin{cases} -K^2 u & r < a \\ +\kappa^2(r) u & a < r < b \\ -k^2(r) u & r > b \end{cases} \tag{14.8.125}$$

where

$$\begin{aligned} K &= \tfrac{\sqrt{2m(E+V_0)}}{\hbar} \qquad & r < a \\ \kappa(r) &= \tfrac{\sqrt{2m[V(r)-E]}}{\hbar} \qquad & a < r < b \\ k(r) &= \tfrac{\sqrt{2m[E-V(r)]}}{\hbar} \qquad & b < r \end{aligned} \tag{14.8.126}$$

and m is the reduced mass of the alpha particle

$$m = \frac{m_\alpha M}{m_\alpha + M} . \tag{14.8.127}$$

Here, M is the mass of the daughter nucleus.

Now using the results of the previous section we have from (14.7.122) that

$$T \approx iS \tag{14.8.128}$$

where

$$S = \exp\left[-\int_a^b \kappa(r')\, dr'\right] \tag{14.8.129}$$

and a, the radius of the nuclear potential, and $b = (2Ze^2)/E$ are the two classical turning points. The integral in the exponential is simply

$$\sigma = \int_a^b \sqrt{\frac{4mZe^2}{\hbar^2 r'} - \frac{2mE}{\hbar^2}}\, dr' . \tag{14.8.130}$$

After defining the "alpha particle Bohr radius"

$$a_\alpha = \frac{\hbar^2}{me^2} \tag{14.8.131}$$

the integral may be rewritten in dimensionless form

$$\sigma = \sqrt{\frac{4Z}{a_\alpha b}} \int_{a/b}^1 dx\, \sqrt{1/x - 1} . \tag{14.8.132}$$

Since $a/b << 1$ the integral involved may be approximated by

$$\int_{a/b}^1 dx\, \sqrt{1/x - 1} = \int_0^1 dx\, \sqrt{1/x - 1} - \int_0^{a/b} dx\, \sqrt{1/x - 1} \tag{14.8.133}$$

the first integral is easily evaluated by letting

$$x = \cos^2\theta$$

and yields the value $\pi/2$. The second integral is approximated by

$$\int_0^{a/b} dx\, \sqrt{1/x - 1} \approx \int_0^{a/b} \frac{dx}{\sqrt{x}} = 2\sqrt{a/b} . \tag{14.8.134}$$

Therefore,

$$\sigma = \sqrt{\frac{4Zb}{a_\alpha}} \left[\pi/2 - 2\sqrt{a/b}\right] . \tag{14.8.135}$$

The transmission amplitude may therefore be written

$$S = \exp\left[-2\pi Z\sqrt{E/E_0} + 4\sqrt{(Za)/a_\alpha}\right] \tag{14.8.136}$$

where we have introduced the characteristic energy

$$E_0 = \frac{e^2}{2a_\alpha} \approx 99.3 \text{ keV} . \tag{14.8.137}$$

The transmission coefficient is given by

$$T = S^2 . \tag{14.8.138}$$

However, to compute the transition rate, and hence the half-life of the alpha emitter we have to multiply the transition probability T by the collision frequency f (the number of times per second that the alpha particle hits the barrier). This is easily estimated by

$$f = \frac{v}{2a} = \frac{p}{2ma} \tag{14.8.139}$$

where v is the velocity and p is the momentum of the alpha particle inside the nucleus. We estimate this from the energy E_α of the emitted alpha particle by assuming that the kinetic energy of the alpha particle inside the nucleus is is equal to the energy of the emitted alpha. This is not quite correct and underestimates the kinetic energy somewhat. However, the dominant behaviour comes from the energy dependence in the transmission coefficient and so we are content to use this approximation. This means that

$$f = \frac{1}{a}\sqrt{E_\alpha/(2m)}\ . \tag{14.8.140}$$

Thus, the decay rate is given by

$$\Gamma = fT = \frac{1}{a}\sqrt{E_\alpha/(2m)}\exp\left[-4\pi Z\sqrt{E/E_0} + 8\sqrt{(Za)/a_\alpha}\right]\ . \tag{14.8.141}$$

The half-life is then given by

$$\tau_{1/2} = \frac{\ln 2}{\Gamma}\ . \tag{14.8.142}$$

To see how remarkably successful this model is we compute the half-lives for the two isotopes mentioned in the introduction. For Thorium (Z=90) the daughter nucleus is radium ($Z = 88$). If we use a nuclear potential with a range of $a = 9.00$ fm then we find that

$$T \approx 1.3 \times 10^{-39}\ . \tag{14.8.143}$$

The collision frequency is

$$f = 7.77 \times 10^{20} \quad \text{collisions per second}\ . \tag{14.8.144}$$

Thus, the decay rate is $\Gamma = fT \approx 1.3 \times 10^{-39} \times 7.77 \times 10^{20} = 1.01 \times 10^{-18}$ s. The corresponding half-life is

$$\tau_{1/2} = \frac{\ln 2}{\Gamma} = 6.9 \times 10^{17} \ \ \text{s} \ = 2.2 \times 10^{10} \ \ \text{years}\ . \tag{14.8.145}$$

If we do the same calculation for Po^{212} we get

$$T \approx 8.2 \times 10^{-13}\ . \tag{14.8.146}$$

and

$$f = 1.15 \times 10^{21} \quad \text{collisions per second} \tag{14.8.147}$$

so that

$$\Gamma = 1.15 \times 10^{21} \times 8.2 \times 10^{-13} = 9.43 \times 10^{7} \ \ \text{per second}\ . \tag{14.8.148}$$

Thus, in this case the half-life is

$$\tau_{1/2} = 7.34 \times 10^{-10}\text{s}\ . \tag{14.8.149}$$

Although our results give only qualitative agreement with the experimental results, they must be considered satisfactory in view of the extreme simplicity of the model used. Furthermore they illustrate how, in spite of the small variation in energy of the emitted alphas, the lifetimes can vary over many orders in magnitude due to the very strong energy dependence of the transmission coefficient T.

14.8.2 Detailed Analysis

For the more detailed analysis we approximate the potential as before, but we analyze the wavefunction in more detail. Although we wind up with an eigenvalue equation for the alpha particles in the nuclear potential well, the resulting energies are *complex* since the resulting states are not stationary but decay in time due to the fact that the alpha particles are able to tunnel out. This is due to the fact that the Hamiltonian with the prescribed boundary conditions (only an outgoing wave outside the nucleus $r > b$ and no incoming wave) is not self-adjoint. The imaginary part of the energy now gives the decay rate. As before, the radial s-wave function is given by

$$\psi(r) = \frac{u(r)}{r} . \tag{14.8.150}$$

Then, $u(r)$ is given by

$$u(r) = N \begin{cases} \sin Kr & r < a \\ \kappa(r)^{-1/2}\left[A \exp\left(\int_a^r \kappa(r')\,dr'\right) \right. \\ \left. + B \exp\left(-\int_a^r \kappa(r')\,dr'\right)\right] & a < r < b \\ k(r)^{-1/2} S \exp i\left(\int_b^r k(r')\,dr' - \frac{\pi}{4}\right) & b < r \end{cases} \tag{14.8.151}$$

where N is an overall normalization factor and we have imposed the condition that for $r > b$ there is only an outgoing wave.

Also, we have already written $u(r)$ in the form of a WKB wavefunction for the region $r > a$. In the case at hand, we can simply match the wavefunction and its derivative at $r = a$ and use the WKB matching conditions at $r = b$. Thus, writing as in the previous section,

$$\sigma = \frac{1}{\hbar} \int_a^b \sqrt{2m[V(r) - E]}\,dr \tag{14.8.152}$$

we find

$$\begin{aligned} A &= \frac{\sqrt{\kappa(a)}}{2}\left[\sin Ka \left(1 + \frac{\kappa'(a)}{\kappa^2(a)}\right) + \frac{K}{\kappa(a)} \cos Ka\right] \\ B &= \frac{\sqrt{\kappa(a)}}{2}\left[\sin Ka \left(1 - \frac{\kappa'(a)}{\kappa^2(a)}\right) - \frac{K}{\kappa(a)} \cos Ka\right] . \end{aligned} \tag{14.8.153}$$

Also,

$$\begin{aligned} 2A &= S\,e^{-\sigma} \\ B &= -iS\,e^{\sigma} . \end{aligned} \tag{14.8.154}$$

Here, using

$$V(r) = \frac{Ze^2}{r}$$

we find that

$$-\frac{\kappa'(a)}{\kappa^2(a)} = -\frac{Z}{(\kappa(a)a)^2\kappa(a)a_\alpha} \ . \tag{14.8.155}$$

If we insert typical values as in the previous section we find that

$$-\frac{\kappa'(a)}{\kappa^2(a)} \leq 0.03 \ . \tag{14.8.156}$$

Thus, we can drop these terms in (14.8.153) and get

$$\begin{aligned} A &= \frac{\sqrt{\kappa(a)}}{2}\left[\sin Ka + \frac{K}{\kappa(a)}\cos Ka\right] \\ B &= \frac{\sqrt{\kappa(a)}}{2}\left[\sin Ka - \frac{K}{\kappa(a)}\cos Ka\right] \ . \end{aligned} \tag{14.8.157}$$

Combining (14.8.157) and (14.8.154) we get the eigenvalue equation

$$\tan Ka + \frac{K}{\kappa(a)} = \frac{i}{2}\,e^{-2\sigma}\left[\tan Ka - \frac{K}{\kappa(a)}\right] \ . \tag{14.8.158}$$

This clearly shows that the eigenvalues are complex. On the other hand, all estimates of σ , as we saw in the heuristic estimates in the previous section, lead to values of $e^{-2\sigma} < 10^{-15}$. Thus, the right hand side in the eigenvalue equation is approximately zero and the eigenvalue equation becomes

$$\tan Ka \approx -\frac{K}{\kappa(a)} \ . \tag{14.8.159}$$

Furthermore, for low-lying states and heavy nuclei (high Coulomb barriers) $K << \kappa(a)$. This means that the real part of the eigenvalues is given by

$$\tan Ka \approx 0 \quad K = \frac{n\pi}{a} \ . \tag{14.8.160}$$

These are identical to the energies of a particle in an infinite square well. Thus, the lowest energies are given by

$$E_n = \frac{\hbar^2\pi^2 n^2}{2ma^2} - V_0 \ . \tag{14.8.161}$$

The estimate of σ also shows that the overall normalization may be computed from the interior region of the nucleus. Thus, we have that

$$\begin{aligned} 1 &\approx |N|^2 4\pi \int_0^a u^2(r)\,dr \\ &= |N|^2 4\pi \int_0^a \sin^2 Kr\,dr \\ &= \frac{2\pi|N|^2}{K}[Ka - \frac{1}{2}\sin(2Ka)] \ . \end{aligned} \tag{14.8.162}$$

This yields, using $Ka = n\pi$, that

$$|N|^2 \approx \frac{1}{2\pi a} . \tag{14.8.163}$$

To compute the decay rate we now compute the probability flux out of the nucleus. The radial component of the current for $r > b$ is given by

$$j = \frac{\hbar}{2im}\left[\psi^* \frac{d\psi}{dr} - \psi \frac{d\psi^*}{dr}\right] = \frac{\hbar |NS|^2}{mr^2} . \tag{14.8.164}$$

Here we have used that $u = r\psi$. The decay rate is now given by computing the flux through a large sphere centred at the nucleus.

$$\Gamma = \int j(r)\, r^2\, d\Omega = 4\pi \frac{\hbar |NS|^2}{m} . \tag{14.8.165}$$

We have already evaluated $|N|^2$. The constant S is most easily evaluated by recalling that $\tan Ka \approx -K/\kappa(a) \approx 0$. Then, $\sin Ka \approx -K/\kappa(a)\ \cos Ka$ and $\cos Ka \approx \pm 1$. Thus, from (14.8.153) we get that

$$B \approx \pm \frac{K}{\sqrt{\kappa(a)}} . \tag{14.8.166}$$

After substituting this result into (14.8.154) we get

$$|S|^2 = |B|^2\, e^{-2\sigma} \approx \frac{K^2}{\kappa(a)}\, e^{-2\sigma} . \tag{14.8.167}$$

Therefore, we have that

$$\Gamma = \frac{\hbar K}{2ma}\, \frac{4K}{\kappa(a)}\, e^{-2\sigma} . \tag{14.8.168}$$

The first factor $(\hbar K)/(2ma)$ is clearly just the collision frequency $p/(2ma)$. The remaining factors are just the transmission coefficient. This is therefore qualitatively similar to the results of our heuristic calculation. However, in this case we also have an expression for K and the energy of the alpha particles.

It is also possible to obtain the same result if we just solve for the imaginary part of the energy and realize that the complex energy E must be of the form

$$E = E_0 - i\Gamma/2 \tag{14.8.169}$$

where E_0 is the real part of the energy. This form results from realizing that the probability density for finding the particle inside the nucleus must decay exponentially with the decay constant Γ. But, the time dependence of the wavefunction is given by

$$\Psi(t) = \psi_E\, e^{-iEt} . \tag{14.8.170}$$

And using the form (14.8.169) above we get that

$$|\Psi(t)|^2 = |\psi_E|^2\, e^{-\Gamma t} . \tag{14.8.171}$$

To see how this works we begin with the eigenvalue equation

$$\frac{\tan Ka + \frac{K}{\kappa(a)}}{\tan Ka - \frac{K}{\kappa(a)}} = \frac{i}{2}e^{-2\sigma} \ . \tag{14.8.172}$$

Clearly the eigenvalue K is now complex and since the real part of K say K_0 is given by

$$\tan K_0 a + \frac{K_0}{\kappa_0(a)} \approx 0 \tag{14.8.173}$$

we can set

$$K = K_0 - i\Delta \tag{14.8.174}$$

where $|\Delta| << K_0$.

In this case, since

$$E = \frac{\hbar^2 K^2}{2m} \approx \frac{\hbar^2 K_0^2}{2m} - i\frac{\hbar^2 K_0 \Delta}{m} = E_0 - i\Gamma/2 \tag{14.8.175}$$

we have that

$$\Gamma = \frac{2\hbar^2 K_0 \Delta}{m} \ . \tag{14.8.176}$$

So, we need to solve for Δ.

Now using (14.8.174) and the fact that $|\Delta| << K_0$ we get

$$\begin{aligned} \tan Ka &\approx \frac{\tan K_0 a - i\Delta a}{1 + i\Delta a \tan K_0 a} \\ &\approx \tan K_0 a - i\Delta a \left(1 + \tan^2 K_0 a\right) \\ &= \tan K_0 a - \frac{i\Delta a}{\cos^2 K_0 a} \\ &\approx \tan K_0 a - i\Delta a \ . \end{aligned} \tag{14.8.177}$$

But,

$$\tan K_0 a \approx -\frac{K_0}{\kappa_0(a)} \ . \tag{14.8.178}$$

Therefore,

$$\tan Ka \approx -\frac{K_0}{\kappa_0(a)} - i\Delta a \ . \tag{14.8.179}$$

So we get

$$\frac{i}{2}e^{-2\sigma} = -i\Delta a \frac{-\kappa_0(a)}{2K_0} \ . \tag{14.8.180}$$

Thus,

$$\Delta = \frac{K_0}{\kappa_0(a)} e^{-2\sigma} \ . \tag{14.8.181}$$

Therefore, using (14.8.176) we get that

$$\Gamma = \frac{2\hbar^2 K_0^2}{ma\,\kappa_0(a)}\, e^{-2\sigma}\,. \tag{14.8.182}$$

This is exactly the same result as obtained before.

14.9 Problems

14.1 Use the trial wavefunction

$$\psi(x) = A\, e^{-\alpha|x|}$$

to estimate the ground state energy of a simple harmonic oscillator.

14.2 Use the trial wavefunction (14.4.30) to evaluate the expectation value of the Hamiltonian (14.4.28).

14.3 Use the connection formulae to solve the tunneling problem for a square barrier with the WKB approximation. This result agrees with the exact result in the text.

14.4 Use a variational approach to find the ground state energy for a particle in the potential

$$V(r) = -V_0 e^{-\alpha r^2}$$

if the particle involved is an electron and $\alpha = 5.29 \times 10^{13}$ cm^{-2} , $V_0 = 20$ eV. A numerical answer is required.

14.5 An electron is in the spherically symmetric potential

$$V(r) = A\, r^2(r^2 - a^2)$$

where

$$a = 2.00 \times 10^{-1}\text{nm}$$

$$a^4 A = 1.90 \text{ eV}.$$

Use Rayleigh Ritz to estimate the ground state energy. A numerical answer is required.

14.6 A particle is in a potential

$$V = \begin{cases} mgz & z > 0 \\ \infty & z = 0 \end{cases} .$$

This corresponds to a perfectly elastic ball bouncing on a floor. Find the WKB solution for the ground state energy and wavefunction of this particle.

14.7 Estimate the ground state energy of H_2^+, an ionized hydrogen molecule.

14.8 Use the WKB approximation to solve for the reflection and transmission coefficients for scattering of a particle with momentum $\hbar k$ from the potential given by

$$V = \begin{cases} V_0\left(1 - \frac{|x|}{a}\right) & |x| < b \\ 0 & |x| > b \end{cases}$$

if

$$V_0 >> \frac{\hbar^2 k^2}{2m} \quad b > a > 0 \ .$$

14.9 Repeat problem 14.8 if

$$V_0 << \frac{\hbar^2 k^2}{2m} \quad b > a > 0 \ .$$

14.10 Use a variational technique to find the ground state energy of a particle of mass m in the potential

$$V = \begin{cases} \infty & x < -a \\ V_0 & -a \le x \le a \\ \infty & x > a \end{cases} \ .$$

Hint: Pay close attention to the boundary conditions that the trial wavefunction has to satisfy.

Bibliography

[14.1] It is again possible to find several of the techniques discussed in this chapter, in the book by Rayleigh [12.1].

[14.2] G.N. Watson, *A Treatise on the Theory of Bessel Functions* - 2nd Edition, Cambridge University Press (1962).
I.S. Gradshteyn and I.M. Ryzhik , *Table of Integrals, Series, and Products*, Academic Press, New York (1965).

[14.3] The WKB approximation was applied to wave propagation by Rayleigh in 1872.
It was then applied by: H. Jeffreys, Proc. London Math. Soc. **23**, 428 (1923).
Later it was simultaneously rediscovered and applied to quantum mechanics by:
L. Brillouin, Comptes Rendus **183**, 24 (1926).
H.A. Kramers, Zeits. f. Physik **39**, 828 (1926).

G. Wentzel,Zeits. f. Physik **38**, 518 (1926).
There have also been many papers devoted to the connection formulae for the WKB solutions. Two representative papers are:
R.E. Langer, Phys. Rev. **51**, 669 (1937).
W.H. Furry, Phys. Rev. **71**, 360 (1947).
See also the lectures by W.H. Furry, pages 34-67 in W.E. Brittin, W.B. Downs and J.Downs (editors) - *Lectures in Theoretical Physics Vol. V* - Summer Institute for Theoretical Physics, University of Colorado, Boulder (1962).

[14.4] G. Gamow, Z. Phys. **51**, 204-212 (1928).

[14.5] For alternate approaches to alpha decay see
B.H. Holstein, Am. J. Phys. **64**, 1061-1071 (1996).

Chapter 15

Time-Dependent Perturbation Theory

15.1 Introduction

In the previous chapters we discussed techniques that are applicable if the perturbation is considered as causing a change in the states of the unperturbed system leading to shifts in the energy. This requires that the perturbation be time-independent.

In the present chapter we are not interested in the modifications of the states of the unperturbed system, rather we are interested in the transitions occurring between levels of the unperturbed system under the influence of the perturbation. This means, we are interested in computing transition rates. In general we therefore have time-dependent perturbations.

15.2 Formal Considerations

For the sake of concreteness consider a Hamiltonian

$$H = H_0 + \lambda V(t) \tag{15.2.1}$$

where the perturbation $V(t)$ is written as explicitly time-dependent. Let $\{|n\rangle\}$ be a complete set of eigenkets of H_0 (in the energy representation) so that

$$H_0|n\rangle = E_n|n\rangle \ . \tag{15.2.2}$$

If λ were zero these would be the stationary states of the system, but due to the perturbation, transitions between these states occur. Now, suppose that at some definite time $t = 0$ the system is known (prepared) to be in the state $|n\rangle$. Then at some later time t the system will be in a state $U(t)|n\rangle$ where $U(t)$ is the evolution operator and is given by solving

$$i\hbar\frac{\partial U(t)}{\partial t} = HU(t) \tag{15.2.3}$$

with

$$U(0) = 1 \, . \tag{15.2.4}$$

The probability of finding the system at time t in the state $|m\rangle$ is then given by

$$P_{n,m} = |\langle m|U(t)|n\rangle|^2 \, . \tag{15.2.5}$$

If $n = m$, this gives the probability that the system has not changed in the time t. Thus the main problem in computing transitions reduces to calculating the amplitudes $\langle m|U(t)|n\rangle$. We, theore, reconsider the equations for $U(t)$ in more detail.

Equation (15.2.3) may be combined with the initial condition (15.2.4) in a single integral equation

$$U(t) = 1 - \frac{i}{\hbar} \int_0^t H(t')U(t')\, dt' \, . \tag{15.2.6}$$

We can also extract the behaviour due to H_0 by going to the interaction picture. Thus, let $U_0(t)$ be the evolution operator corresponding to H_0, that is

$$i\hbar \frac{\partial U_0(t)}{\partial t} = H_0 U_0(t) \tag{15.2.7}$$

$$U_0(0) = 1 \, . \tag{15.2.8}$$

We then look for an operator $U_I(t)$ such that

$$U(t) = U_0(t) U_I(t) \tag{15.2.9}$$

and since U_0 is unitary we have

$$U_I(t) = U_0^\dagger(t) U(t) \, . \tag{15.2.10}$$

Substituting this into (15.2.3) and (15.2.4), and using (15.2.7) and (15.2.8) we obtain the equation for $U_I(t)$.

$$i\hbar \frac{\partial U_I(t)}{\partial t} = \lambda H_I U_I(t) \tag{15.2.11}$$

with

$$U_I(0) = 1 \, . \tag{15.2.12}$$

$$H_I(t) = U_0^\dagger(t) V(t) U_0(t) \, . \tag{15.2.13}$$

This again gives rise to an equivalent integral equation

$$U_I(t) = 1 - \frac{i\lambda}{\hbar} \int_0^t H_I(t') U_I(t')\, dt' \, . \tag{15.2.14}$$

A formal solution, which is also the basis of an approximation procedure, is now obtained by iterating (15.2.14). Doing this, we get

$$\begin{aligned} U_I(t) &= 1+\left(\frac{\lambda}{i\hbar}\right)\int_0^t H_I(t')\,dt' \\ &+ \left(\frac{\lambda}{i\hbar}\right)^2\int_0^t dt'\int_0^{t'} H_I(t')H_I(t'')U_I(t'')\,dt'' \end{aligned} \tag{15.2.15}$$

or continuing the procedure we arrive at:

$$U_I(t) = 1+\sum_{n=1}^{\infty} U_I^{(n)}(t) \tag{15.2.16}$$

where

$$\begin{aligned} U_I^{(n)}(t) &= \left(\frac{\lambda}{i\hbar}\right)^n \int_0^t dt_1 \int_0^{t_1} dt_2 \ldots \\ & \int_0^{t_{n-1}} dt_n\, H_I(t_1)H_I(t_2)\ldots H_I(t_n)\,. \end{aligned} \tag{15.2.17}$$

It is important to note the order of the operators $H_I(t)$ since they need not commute for different times. Combining (15.2.17) with the definition of U_I in terms of $U_0 \equiv U^{(0)}$ we get

$$U(t) = U^{(0)}(t)+\sum_{n=1}^{\infty} U^{(n)}(t) \tag{15.2.18}$$

where

$$\begin{aligned} U^{(n)}(t) &= U^{(0)}(t)\left(\frac{\lambda}{i\hbar}\right)^n \int_0^t dt_1 \int_0^{t_1} dt_2 \ldots \\ & \int_0^{t_{n-1}} dt_n\, U^{(0)\dagger}(t_1)V(t_1)U^{(0)}(t_1) \\ \times\ & U^{(0)\dagger}(t_2)V(t_2)U^{(0)}(t_2)\ldots U^{(0)\dagger}(t_n)V(t_n)U^{(0)}(t_n)\,. \end{aligned} \tag{15.2.19}$$

If we now use the fact that H_0 is time-independent we have that

$$U_0(t) = U^{(0)}(t) = \exp\left(-\frac{iH_0 t}{\hbar}\right)\,. \tag{15.2.20}$$

In that case the transition probability, to different orders in λ , can be written explicitly in terms of the transition amplitudes

$$\langle m'|U(t)|m\rangle = \sum_{n=1}^{\infty}\langle m'|U^{(n)}(t)|m\rangle \tag{15.2.21}$$

where $U^{(n)}(t)$ is given by (15.2.19). Writing out the first few orders we have

$$\langle m'|U^{(0)}(t)|m\rangle = \exp\left(-\frac{iE_{m'}t}{\hbar}\right)\delta_{m,m'} \tag{15.2.22}$$

$$\begin{aligned}&\langle m'|U^{(1)}(t)|m\rangle\\ =\;& \frac{\lambda}{i\hbar}\int_0^t dt_1 e^{-i(E_{m'}/\hbar)(t-t_1)}\langle m'|V(t_1)|m\rangle e^{-i(E_m/\hbar)t_1}\end{aligned} \tag{15.2.23}$$

$$\begin{aligned}\langle m'|U^{(2)}(t)|m\rangle = \left(\frac{\lambda}{i\hbar}\right)^2\sum_n\int_0^t dt_1\int_0^{t_1} dt_2 e^{-i(E_{m'}/\hbar)(t-t_1)}\times\\ \times\langle m'|V(t_1)|n\rangle e^{-i(E_m/\hbar)(t_1-t_2)}\langle n|V(t_2)|m\rangle e^{-i(E_m/\hbar)t_2}\end{aligned} \tag{15.2.24}$$

where the sum over n runs over the complete set of kets $|n\rangle$.

15.3 Transition Amplitudes

This whole procedure discussed in the previous section may also be obtained in a much less formal and more direct manner which we now describe. We start with the exact Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = H|\Psi(t)\rangle \tag{15.3.25}$$

and expand $|\Psi(t)\rangle$ in the terms of the stationary states of H_0. This is the same as going to the interaction picture. Then we get

$$|\Psi(t)\rangle = \sum_n a_n(t)e^{-iE_nt/\hbar}|n\rangle\ . \tag{15.3.26}$$

Note that the coefficients $a_n(t)$ have to be explicitly time dependent. Substituting this in the Schrödinger equation (15.3.25) we obtain

$$\begin{aligned}i\hbar\;&\sum_n\frac{da_n}{dt}e^{-iE_nt/\hbar}|n\rangle+\sum_n a_nE_ne^{-iE_nt/\hbar}|n\rangle\\ =\;&\sum_n a_n[H_0+\lambda V]e^{-iE_nt/\hbar}|n\rangle\\ =\;&\sum_n a_n\lambda Ve^{-iE_nt/\hbar}|n\rangle+\sum_n a_nE_ne^{-iE_nt/\hbar}|n\rangle\ .\end{aligned} \tag{15.3.27}$$

If we now form the inner product with $\langle k|$ and call

$$E_k - E_n = \hbar\omega_{kn} \tag{15.3.28}$$

we get

$$\frac{da_k}{dt} = \frac{\lambda}{i\hbar}\sum_n\langle k|V|n\rangle a_n\, e^{i\omega_{kn}t}\ . \tag{15.3.29}$$

This differential equation is exact and corresponds to the equation (15.2.11) for the evolution operator in the interaction picture, namely

$$i\hbar\frac{dU_I}{dt} = H_IU_I(t)\ . \tag{15.3.30}$$

If furthermore in the past, say for $t < 0$, $V(t)$ vanished and at that time the system was in the state $|m\rangle$ then we have

$$\frac{da_k}{dt} = 0 \quad \text{for} \quad t < 0 \tag{15.3.31}$$

so that a_k = constant. This constant is either 0 or 1 depending on whether $k \neq m$ or $k = m$. Thus, we have an initial condition for (15.3.29), namely

$$a_k(0) = \delta_{k,m} \,. \tag{15.3.32}$$

We can now rewrite (15.3.29) as an integral equation incorporating (15.3.32). To emphasize the fact that the initial state was $|m\rangle$ we write $a_{k,m}$ instead of a_k. We then have

$$a_{k,m}(t) = \delta_{k,m} + \frac{\lambda}{i\hbar}\sum_n \int_0^t dt' \, \langle k|V(t')|n\rangle a_{n,m}(t')\, e^{i\omega_{kn}t'} \,. \tag{15.3.33}$$

This is analogous to equation (15.2.14). It is now a simple matter to iterate this equation to obtain equations analogous to (15.2.22) and (15.2.23) etc. Thus, the zeroth order term in λ is

$$a^{(0)}_{k,m}(t) = \delta_{k,m} \tag{15.3.34}$$

and the first order term is

$$a^{(1)}_{k,m}(t) = \frac{\lambda}{i\hbar}\int_0^t dt' \, \langle k|V(t')|m\rangle\, e^{i\omega_{km}t'} \,. \tag{15.3.35}$$

The second order term is obtained by inserting this result in equation (15.3.33). We now apply this result to several special cases which are of particular interest in practice.

15.4 Time-Independent Perturbation

We begin with equation (15.3.35) and consider transitions during the time from $-T$ to T. Then the probability amplitude to find a particle in the state $|k\rangle$ at a time $t > T$ if it was in the state $|m\rangle$ at time $-T$ is given to first order by

$$a^{(1)}_{k,m}(t) = \frac{\lambda}{i\hbar}\int_{-T}^{T} dt' \, \langle k|V(t')|m\rangle\, e^{i\omega_{km}t'} \,. \tag{15.4.36}$$

The transition probability per unit time w to make a transition from an *initial state* $|m\rangle$ to a *single final state* $|n\rangle$ is therefore given by

$$\begin{aligned} w^{\text{single}}_{f,i} &= \frac{|a_{n,m}(T)|^2}{2T} \\ &= \left(\frac{\lambda}{\hbar}\right)^2 |\langle n|V|m\rangle|^2 \int_{-T}^{T} e^{i\omega_{nm}t'}\, dt' \, \frac{1}{2T}\int_{-T}^{T} e^{-i\omega_{nm}t''}\, dt'' \,. \end{aligned} \tag{15.4.37}$$

In general one wants to calculate this rate for very long times T so that the result does not depend on T. To do this we take the limit as $T \to \infty$. Then, we find

$$
\begin{aligned}
& \lim_{T\to\infty} \int_{-T}^{T} e^{i\omega_{nm}t'}\,dt'\,\frac{1}{2T}\int_{-T}^{T} e^{-i\omega_{nm}t''}\,dt'' \\
= \; & \lim_{T\to\infty} 2\pi\delta(\omega_{n,m})\,\frac{1}{2T}\int_{-T}^{T} e^{-i\omega_{nm}t''}\,dt'' \\
= \; & \lim_{T\to\infty} 2\pi\delta(\omega_{n,m})\,\frac{1}{2T}\int_{-T}^{T} dt'' \\
= \; & 2\pi\delta\left(\frac{E_n - E_m}{\hbar}\right) \\
= \; & 2\pi\hbar\,\delta\left(E_n - E_m\right) \; .
\end{aligned}
\qquad (15.4.38)
$$

Combining this with equation (15.4.37) we have

$$
w_{f,i}^{\text{single}} = \frac{2\pi}{\hbar}|\lambda\langle n|V|m\rangle|^2\delta\left(E_n - E_m\right) \; . \qquad (15.4.39)
$$

Clearly, the delta-function expresses the conservation of energy and such a transition can only occur between two levels with the same energy. This means we have to have either degenerate levels or, what is more interesting, elastic scattering. In practice it is this latter case that is of interest. In such a situation one cannot measure the exact momentum of the final state since the detector has a finite resolution so that the best one can do is to state that the momentum after scattering (the momentum of the final state) lies between $\hbar k$ and $\hbar(k+dk)$. This situation is then described by introducing a "density of final states" $\rho(k)$ which tells us how many states lie between k and $k + dk$. In terms of $\rho(k)$ this number is given by

$$
\rho(k)\,d^3k = \rho(k)\,k^2 dk\,d\Omega \; . \qquad (15.4.40)
$$

The density of final states that also conserves energy is then given by (notice that now the function ρ is defined by its argument)

$$
\rho(E_k) = \delta(E_k - E_m)\rho(k) \; . \qquad (15.4.41)
$$

The transition probability per unit time to such an interval of states is therefore given by

$$
w_{k,m} = w_{f,i}^{\text{single}}\rho(k) = \frac{2\pi}{\hbar}|\lambda\langle n|V|m\rangle|^2\rho(E_k) \; . \qquad (15.4.42)
$$

This is the celebrated *Fermi's Golden Rule.* We now rederive it for a time-dependent perturbation.

15.5 Periodic Perturbation of Finite Duration

The perturbation part of the Hamiltonian is in this case assumed to have the special form

$$\langle k|V(t)|m\rangle = \begin{cases} 0 & \text{for } t<0\,,\ t>T \\ 2\langle k|v|m\rangle \sin\omega t & \text{for } 0<t<T \end{cases} \quad . \tag{15.5.43}$$

Here v is assumed to be some time-independent operator. The special form of the time dependence during the interval $0<t<T$ is not a restriction. We could have taken an arbitrary time dependence in the interval $0<t<T$. In that case $\langle k|v|m\rangle$ would be just the appropriate Fourier component for the frequency ω.

We can immediately compute the first order term $a_k^{(1)}(t)$. For $t \geq T$ we then have as before

$$a_k^{(1)}(t) = \frac{\lambda}{i\hbar}\langle k|v|m\rangle \int_0^T 2\sin\omega t'\, e^{i\omega_{km}t'}\, dt' \tag{15.5.44}$$

or

$$a_k^{(1)}(t) = \frac{\lambda}{i\hbar}\langle k|v|m\rangle \left[\frac{e^{i(\omega_{km}-\omega)T}-1}{\omega_{km}-\omega} - \frac{e^{i(\omega_{km}+\omega)T}-1}{\omega_{km}+\omega}\right] . \tag{15.5.45}$$

Since $|\langle k|v|m\rangle|$ is assumed small, the transition probability will be appreciable only if one of the denominators almost vanishes. This leads to two cases.
Case 1

$$\omega_{km} - \omega \approx 0 \tag{15.5.46}$$

or

$$E_k \approx E_m + \hbar\omega \quad . \tag{15.5.47}$$

Case 2

$$\omega_{km} + \omega \approx 0 \tag{15.5.48}$$

or

$$E_k \approx E_m - \hbar\omega \quad . \tag{15.5.49}$$

Thus, we may interpret that the effect of the perturbation is to transmit to or absorb from the system one quantum $\hbar\omega$ of energy. For a perturbation with only one sharp frequency, as given, the probability is appreciable only if one of the "resonance" conditions (15.5.47) or (15.5.49) is satisfied.

We now specialize this problem even further and consider the transition probability from a bound state $|m\rangle$ to a continuum state $|k\rangle$. In that case $E_k > E_m$ and only case 1 applies. The probability of finding the system in the state $|k\rangle$ if it was initially in the state $|m\rangle$ is then to lowest order in λ given by

$$|a_k^{(1)}(t)|^2 = \left|\frac{\lambda}{\hbar}\langle k|V|m\rangle\right|^2 \cdot \left|\frac{e^{i(\omega_{km}-\omega)T}-1}{\omega_{km}-\omega}\right|^2 \tag{15.5.50}$$

or

$$|a_k^{(1)}(t)|^2 = \frac{4\lambda^2|\langle k|V|m\rangle|^2}{\hbar^2(\omega_{km}-\omega)^2}\sin^2[(\omega_{km}-\omega)T/2] \ . \tag{15.5.51}$$

Intuitively one would expect this probability to be proportional to T, the length of time the interaction was on. If one considers the function

$$\frac{\sin^2[(\omega_{km}-\omega)T/2]}{(\omega_{km}-\omega)^2} = \frac{\sin^2 x}{4x^2}T^2 \tag{15.5.52}$$

where

$$x = (\omega_{km}-\omega)T/2 \tag{15.5.53}$$

we see that for reasonably small T and due to the resonance condition $x \approx 0$ that $|a_k^{(1)}(t)|^2$ is essentially proportional to T^2. This seems to contradict our intuition since we would expect the number of quanta inducing the transition to be proportional to the length of time the interaction is on. The answer lies, in fact, in the finite time over which the interaction is effective. Since a sinusoidal signal lasting a time T can be thought of as a pulse, of length T, such a signal is not monochromatic with frequency ω, but has its energy distributed over a band width proportional to $1/T$. In particular the Fourier transform $F(\omega')$ of the function

$$f(t) = \begin{cases} 0 & \text{for} \quad t \le 0\,,\, t > T \\ \sin\omega t & \text{for} \quad 0 \le t \le T \end{cases} \tag{15.5.54}$$

is

$$F(\omega') = \frac{1}{2\pi i}\left[e^{i(\omega'+\omega)T/2}\frac{\sin(\omega'+\omega)T/2}{(\omega'+\omega)} - e^{i(\omega'-\omega)T/2}\frac{\sin(\omega'-\omega)T/2}{(\omega'-\omega)}\right] . \tag{15.5.55}$$

For the region of interest to us, namely $\omega' > 0$, $\omega > 0$ this function peaks at $\omega' = \omega$ and simplifies to

$$F(\omega') \approx \frac{i}{2\pi}e^{i(\omega'-\omega)T/2}\frac{\sin(\omega'-\omega)T/2}{(\omega'-\omega)} \ . \tag{15.5.56}$$

The points ω'' at which

$$|F(\omega'')| = \frac{1}{2}|F(\omega)| \tag{15.5.57}$$

determine the "width" of this function. They are approximately given by

$$\omega'' = \omega \pm \frac{2}{T}\times 1.9 \ . \tag{15.5.58}$$

Thus,

$$\Delta\omega' = 2\times\frac{2}{T}\times 1.9 = \frac{7.6}{T} \tag{15.5.59}$$

as stated. This shows why, for a unit frequency interval centred about the peak of this spectral distribution, the energy of this pulse is proportional to the square of the duration of the pulse.

In practice, as already discussed above, one does not usually have such short pulses of radiation and one does not usually measure the transition to exactly one final state $|k\rangle$ but rather to a set of such states all with approximately the same energy and hence all approximately satisfying the "resonance" condition. In such a case one is interested in the rate at which such transitions occur. Thus, one calculates w, the transition probability per unit time. This is given to lowest order in λ by

$$w = \frac{1}{T}\int |a_k^{(1)}(t)|^2 \rho(k) dE_k \tag{15.5.60}$$

where $\rho(k)dE_k$ is the number of states per unit volume with energy between E_k and $E_k + dE_k$. Changing the variable of integration from E_k to

$$x = \frac{1}{2\hbar}(E_k - E_m - \hbar\omega)T \tag{15.5.61}$$

so that

$$dx = \frac{T}{2\hbar} dE_k\ , \tag{15.5.62}$$

and substituting from (15.5.53) we get

$$w = \frac{2\lambda^2}{\hbar}\int |\langle k|v|m\rangle|^2 \rho(k) \frac{\sin^2 x}{x^2}\, dx\ . \tag{15.5.63}$$

Now in order to evaluate this integral we make some further approximations.

If T becomes large, the function $\sin^2 x/x^2$ has a high narrow peak with respect to the k variable. Thus, the density of final states $\rho(k)$ and the matrix element $\langle k|v|m\rangle$ are essentially constant over the range where this function is large and may be taken outside the integral. Furthermore for very large T we may extend the limits of integration from $-\infty$ to ∞. Thus,

$$w \approx \frac{2}{\hbar}|\lambda\langle k|v|m\rangle|^2 \rho(k)\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2}\, dx \tag{15.5.64}$$

so that to a good approximation

$$w = \frac{2\pi}{\hbar}|\lambda\langle k|v|m\rangle|^2 \rho(k)\ . \tag{15.5.65}$$

This formula is again the celebrated *Fermi's Golden Rule*. It crops up over and over again in many problems. To illustrate the use of this first-order theory we now apply it to a physically interesting problem.

15.6 Photo-Ionization of Hydrogen Atom

The problem we are interested in consists of calculating the transition probability per unit time that a hydrogen atom in its ground state placed in a high frequency electromagnetic field ejects an electron into a solid angle lying between Ω and $\Omega + d\Omega$.

The Hamiltonian for the electron, assuming it is bound in a fixed Coulomb potential, is then given by

$$H = H_0 + V \tag{15.6.66}$$

where H_0 is the hydrogen atom Hamiltonian

$$H_0 = \frac{\mathbf{p}^2}{2m} - \frac{e^2}{r} \tag{15.6.67}$$

and

$$V = e\mathcal{E}\mathbf{r}\cdot\hat{\epsilon}\, 2\sin\omega t\ . \tag{15.6.68}$$

This is called the dipole approximation. Here $\mathcal{E}$ has the obvious interpretation as the strength of the applied electric field and $\hat{\epsilon}$ is a unit vector in the direction of the electric field (the polarization). The initial state ψ_{in} of the electron is the ground state of an electron in the hydrogen atom

$$\psi_{in} = \psi_{1,0,0}(r) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0}\ . \tag{15.6.69}$$

The final state ψ_{out} is a positive energy (scattering state) of an electron in the presence of a point charge. The resulting wave function is very complicated due to the very long range of the Coulomb potential. We, theore, assume that the Coulomb potential is effectively screened by surrounding matter so that the ejected electron is free. [1] The final electron state is then well approximated by a plane wave

$$\psi_{out} = (2\pi)^{-3/2}\, e^{i\mathbf{k}\cdot\mathbf{r}}\ . \tag{15.6.70}$$

We also need to calculate the density of final states. For this purpose it is simpler first to consider the electron placed in a very large cube of sides L and then later take the limit $L \to \infty$. At the surface of the box we impose periodic boundary conditions since these are the simplest boundary conditions yielding a self-adjoint extension of the momentum operator. In that case we have

$$\psi_{out} = (L)^{-3/2}\, e^{i\mathbf{k}\cdot\mathbf{r}} \tag{15.6.71}$$

with

$$\mathbf{k} = \frac{2\pi}{L}(n_x, n_y, n_z) \tag{15.6.72}$$

---

[1] In interplanetary space, which is the best vacuum known, there are only about 10 free protons and electrons per cm^3. This, however, suffices to already shield the Coulomb field for distances greater than 1 m.

where (n_x, n_y, n_z) are integers. The counting now proceeds just as for the black body radiation (section 1.2). Thus the number of modes in the range between k_x and $k_x + dk_x$ etc. is

$$\begin{aligned} dN &= \Delta n_x \Delta n_y \Delta n_z = \frac{L}{2\pi} dk_x \frac{L}{2\pi} dk_y \frac{L}{2\pi} dk_z \\ &= \left(\frac{L}{2\pi}\right)^3 d^3k = \left(\frac{L}{2\pi}\right)^3 k^2\, dk\, \sin\theta\, d\theta\, d\varphi \ . \end{aligned} \qquad (15.6.73)$$

The use of spherical coordinates in k-space is convenient since energy conservation fixes $|\mathbf{k}| = k$ according to

$$E_k = E_0 + \hbar\omega \qquad (15.6.74)$$

or

$$\frac{(\hbar k)^2}{2m} = -\frac{1}{2} mc^2 \alpha^2 + \hbar\omega \ . \qquad (15.6.75)$$

From equation (15.6.73) we get that

$$\rho(k)\, dE_k = \frac{1}{L^3} dN = \frac{1}{(2\pi)^3} k^2\, dk\, \sin\theta\, d\theta\, d\varphi \ . \qquad (15.6.76)$$

However, we also have

$$dE_k = \frac{\hbar^2}{2m} 2k dk \qquad (15.6.77)$$

and hence

$$\rho(k) d\Omega = \frac{m}{(8\pi)^3 \hbar^2}\, k\, \sin\theta\, d\theta\, d\varphi \ . \qquad (15.6.78)$$

Next we need to evaluate the matrix element

$$\begin{aligned} \langle \mathbf{k} | v | 1,0,0 \rangle &= \int \psi^*_{out}(\mathbf{r}) e\mathcal{E}\hat{\epsilon}\cdot\mathbf{r}\ \psi_{in}(\mathbf{r})\, d^3r \\ &= \frac{e\mathcal{E}}{\sqrt{\pi a_0^3}} \frac{1}{\sqrt{(2\pi)^3}} \int e^{-i\mathbf{k}\cdot\mathbf{r}}\, \hat{\epsilon}\cdot\mathbf{r}\, e^{-r/a_0}\, d^3r \ . \end{aligned} \qquad (15.6.79)$$

The integral may be written as $\hat{\epsilon}\cdot\mathbf{I}$ where

$$\mathbf{I} = \int e^{-i\mathbf{k}\cdot\mathbf{r}}\, \mathbf{r}\, e^{-r/a_0}\, d^3r \ . \qquad (15.6.80)$$

Such vector integrals may be more easily evaluated by using the following symmetry argument. The direction of **I** must be the same as **k** since that is the only vector occuring in the integrand besides the integration variable itself. We can therefore write

$$\mathbf{I} = \mathbf{k} I \qquad (15.6.81)$$

and take the inner product of this equation with $\mathbf{k}$ to get

$$k^2 I = \int e^{-i\mathbf{k}\cdot\mathbf{r}}\,\mathbf{k}\cdot\mathbf{r}\,e^{-r/a_0}\,d^3r\,. \tag{15.6.82}$$

This integral is easier to evaluate than the original expression (15.6.79) since it involves only the angle between $\mathbf{k}$ and $\mathbf{r}$, rather than two angles.

We now choose the z-axis in the integrand to be parallel to $\mathbf{k}$. Then,

$$k^2 I = \int_0^\infty r^2 dr\,e^{-r/a_0}\int_0^\pi e^{-ikr\cos\theta'}\,kr\cos\theta'\sin\theta' d\theta'\int_0^{2\pi} d\varphi' \tag{15.6.83}$$

where $\mathbf{k}\cdot\mathbf{r} = kr\cos\theta'$. The φ' integration is immediate and yields 2π while the θ' integration is easily performed by introducing $u = \cos\theta'$. Then we get

$$k^2 I = 2\pi k\int_0^\infty r^3 dr\,e^{-r/a_0}\int_{-1}^1 e^{-ikru}\,u\,du \tag{15.6.84}$$

These integrals are standard and yield

$$k^2 I = \frac{32\pi i e\mathcal{E}a_0^3(ka_0)^2}{[1+(ka_0)^2]^3}\ . \tag{15.6.85}$$

Hence, we have the matrix element

$$\langle\mathbf{k}|v|1,0,0\rangle = \frac{32\pi i e\mathcal{E}a_0^5\hat{\epsilon}\cdot\mathbf{k}}{(2\pi)^{3/2}(\pi a_0^3)^{1/2}\,[1+(ka_0)^2]^3}\ . \tag{15.6.86}$$

Combining this with our expression for $\rho(k)$, the density of final states, we can use Fermi's golden rule to get

$$dw = \frac{32(e\mathcal{E}a_0^2)^2 m}{\pi^4\hbar^3}\frac{(ka_0)^3}{[1+(ka_0)^2]^6}\cos^2\theta\sin\theta\,d\theta\,d\varphi \tag{15.6.87}$$

where

$$k\cos\theta = \hat{\epsilon}\cdot\mathbf{k}\ . \tag{15.6.88}$$

With $d\Omega = \sin\theta\,d\theta\,d\varphi$ we see that $dw/d\Omega \propto \cos^2\theta$. This simply means that the electron is most likely to be ejected in the direction in which the incident photons are polarized. Also the energy for which $dw/d\Omega$ is maximum is given by setting $d^2w/d\Omega dk = 0$ and solving for k. The result is

$$k = \frac{1}{\sqrt{3}\,a_0} = \frac{me^2}{\sqrt{3}\,\hbar^2}\,. \tag{15.6.89}$$

Hence the energy is

$$E = \frac{(\hbar k)^2}{2m} = \frac{1}{6}mc^2\alpha^2\,. \tag{15.6.90}$$

This is in fact one third times the ground state energy of the hydrogen atom.

We next consider two other types of approximations that are frequently used in practice.

15.7 The Adiabatic Approximation

The perturbation method developed above is generally valid if $\lambda V(t)$ is small for all t, the criterion of smallness being the same as for time-independent perturbation theory. It is also possible to develop approximation methods based on how rapidly $V(t)$ varies with time. Thus, we have an approximation based on very slow time variation, called the adiabatic approximation, which we now discuss and we also have an approximation based on very rapid time variation, called the sudden approximation which we discuss in the next section.

In the adiabatic approximation $|\lambda \partial V/\partial t|$ is assumed to be small, so we try to arrange the computation in such a manner that this term will appear. We start as always with

$$H(t) = H_0 + \lambda V(t) \tag{15.7.91}$$

and using the slow variation of $H(t)$ write the "instantaneous" Schrödinger equation

$$H(t)u_n(t) = E_n(t)u_n(t) \ . \tag{15.7.92}$$

In this equation t is treated simply as a parameter. What we are tacitly assuming in writing (15.7.92) is that if at $t = 0$ the system is in the state $u_n(0)$ with energy $E_n(0)$ then at a slightly later time the system will be in the state approximated by

$$u_n(t) \exp\left(-\frac{i}{\hbar}\int_0^t E_n(t')\, dt'\right) \tag{15.7.93}$$

with energy $E_n(t)$ where the energy $E_n(t)$ varies very slowly. Thus, if we write the time dependent Schrödinger equation

$$i\hbar\frac{\partial \Psi}{\partial t} = H(t)\Psi \tag{15.7.94}$$

and expand Ψ in terms of the slowly varying "eigenstates" $u_n(t)$ so that

$$\Psi = \sum_k a_k(t)u_k(t) \exp\left(-\frac{i}{\hbar}\int_0^t E_k(t')\, dt'\right) \ . \tag{15.7.95}$$

Then the coefficients $a_n(t)$ should be approximately constant and equal to $\delta_{k,n}$. We further assume that the $u_n(t)$ are orthonormal so that

$$(u_n(t), u_k(t)) = \delta_{n,k} \ . \tag{15.7.96}$$

They are of course complete since $H(t)$ is assumed self- adjoint. Substituting (15.7.95) into (15.7.94) yields

$$\begin{aligned} i\hbar\sum_k \left[\dot{a}_k u_k + a_k \dot{u}_k - \frac{i}{\hbar}E_k a_k\right] \exp\left(-\frac{i}{\hbar}\int_0^t E_k(t')\, dt'\right) \\ = \sum_k a_k H u_k \exp\left(-\frac{i}{\hbar}\int_0^t E_k(t')\, dt'\right) \ . \end{aligned} \tag{15.7.97}$$

Using (15.7.92) then yields

$$\sum_k [\dot{a}_k u_k + a_k \dot{u}_k] \exp\left(-\frac{i}{\hbar}\int_0^t E_k(t')\,dt'\right) = 0 \ . \tag{15.7.98}$$

Taking the inner product of this equation with

$$u_m \exp\left(-\frac{i}{\hbar}\int_0^t E_m(t')\,dt'\right) \tag{15.7.99}$$

we obtain the following equation which is still exact

$$\dot{a}_m = \sum_k a_k (u_m, \dot{u}_k) \exp\left(-i\int_0^t \omega_{k,m}(t')\,dt'\right) \ . \tag{15.7.100}$$

Here we have defined

$$\hbar\omega_{k,m}(t) = E_k(t) - E_m(t) \ . \tag{15.7.101}$$

In order to solve (15.7.100) we need to compute $(u_m, \partial u_k/\partial t)$. It is at this point that the slow variation of $V(t)$ appears. From (15.7.92) we get by differentiation

$$\frac{\partial H}{\partial t} u_k + H\frac{\partial u_k}{\partial t} = \frac{\partial E_k}{\partial t} u_k + E_k \frac{\partial u_k}{\partial t} \ . \tag{15.7.102}$$

Taking the inner product with u_m we find

$$\left(u_m, \frac{\partial H}{\partial t} u_k\right) + E_m\left(u_m, \frac{\partial u_k}{\partial t}\right) = \frac{\partial E_k}{\partial t}(u_m, u_k) + E_k\left(u_m, \frac{\partial u_k}{\partial t}\right) \tag{15.7.103}$$

and hence for $m \neq k$

$$\left(u_m, \frac{\partial u_k}{\partial t}\right)(E_k - E_m) = \left(u_m, \frac{\partial H}{\partial t} u_k\right) \tag{15.7.104}$$

or more compactly,

$$\left(u_m, \frac{\partial u_k}{\partial t}\right) = \frac{\lambda}{\hbar\omega_{k,m}}\left(u_m, \frac{\partial V}{\partial t} u_k\right) \quad k \neq m \ . \tag{15.7.105}$$

We also need $(u_m, \partial u_m/\partial t)$. From $(u_m, u_m) = 1$ we get

$$\left(u_m, \frac{\partial u_m}{\partial t}\right) + \left(\frac{\partial u_m}{\partial t}, u_m\right) = 0 \tag{15.7.106}$$

or

$$\left(u_m, \frac{\partial u_m}{\partial t}\right) + \left(u_m, \frac{\partial u_m}{\partial t}\right)^* = 0 \ . \tag{15.7.107}$$

Thus,

$$\left(u_m, \frac{\partial u_m}{\partial t}\right) = i\alpha(t) \tag{15.7.108}$$

where α is real. If we now consider changing the phase of $u_m(t)$ by an amount $\gamma(t)$ to

$$u'_m(t) = u_m(t)e^{i\gamma(t)} \ . \tag{15.7.109}$$

Then,

$$\begin{aligned}\left(u'_m, \frac{\partial u'_m}{\partial t}\right) &= \left(u_m, \frac{\partial u_m}{\partial t}\right) + i\frac{\partial \gamma(t)}{\partial t} \\ &= i\left(\alpha(t) + \frac{\partial \gamma}{\partial t}\right) \ .\end{aligned} \tag{15.7.110}$$

Therefore by choosing

$$\gamma = -\int_0^t \alpha(t')\,dt' \tag{15.7.111}$$

so that

$$\alpha(t) + \frac{\partial \gamma}{\partial t} = 0 \tag{15.7.112}$$

we get that

$$\left(u'_m, \frac{\partial u'_m}{\partial t}\right) = 0 \ . \tag{15.7.113}$$

We henceforth assume that the phase of u_m has been chosen in this way and can be ignored. We therefore drop the primes. This dropping of the phase, as we shall see in section 15.10, is not always possible and can then lead to a very interesting result. However, for the time being we assume that this phase is of no physical significance. Then we have

$$\left(u_m, \frac{\partial u_m}{\partial t}\right) = 0 \ . \tag{15.7.114}$$

In that case we arrive at the following exact equation replacing the Schrödinger equation (15.7.94).

$$\frac{da_m}{dt} = -\sum_k \frac{\lambda}{\hbar\omega_{k,m}}\left(u_m, \frac{\partial V(t)}{\partial t}u_k\right) e^{-i\int_0^t \omega_{k,m}(t')\,dt'}\, a_k \ . \tag{15.7.115}$$

Our approximation now consists of recalling that $V(t)$ and hence $E_n(t)$ and $u_n(t)$ vary slowly with t. Thus, as a first approximation for a_k we choose these quantities constant. As initial condition we further assume that at $t = 0$ the system is in the state u_n so that

$$a_k(t) = a_k(0) = \delta_{k,n} \ . \tag{15.7.116}$$

With the approximations above (15.7.115) becomes

$$\frac{da_m}{dt} = -\frac{\lambda}{\hbar\omega_{n,m}}\left(u_m, \frac{\partial V(t)}{\partial t}u_n\right) e^{-i\omega_{n,m}t} \quad m \neq n \ . \tag{15.7.117}$$

This combined with the initial condition yields

$$a_m(t) = \frac{\lambda}{i\hbar\omega_{n,m}^2}\left(u_m, \frac{\partial V(t)}{\partial t}u_n\right)\left[e^{-i\omega_{n,m}t} - 1\right] \quad m \neq n \ . \tag{15.7.118}$$

This is then the solution in the case of the adiabatic approximation. It follows from this equation that for $m \neq n$, a_m oscillates and does not increase monotonely with t. In fact

$$|a_m(t)| = \left|\frac{\lambda\frac{T}{2\pi}\left(u_m, \frac{\partial V(t)}{\partial t}u_n\right)}{E_n - E_m}\right| 2|\sin(\omega_{n,m}t/2)| \tag{15.7.119}$$

where $T = 2\pi/\omega_{n,m}$. However if $V(t)$ is itself oscillatory with a frequency ω comparable to $\omega_{n,m}$ then $V(t)$ can no longer be considered to vary slowly and, in fact, the approximation breaks down since we get "resonance". In this case small changes in V can cause large changes in $|a_m(t)|$ as we saw in the perturbation treatment. This means we can no longer neglect $\partial V/\partial t$. To be specific, suppose we have

$$H = H_0 + 2V_0 \sin\omega t \ . \tag{15.7.120}$$

Thus,

$$\frac{\partial V}{\partial t} = 2\omega V_0 \cos\omega t \ . \tag{15.7.121}$$

We assume that both $T\ \partial V/\partial t$ and V are small so that $a_n(t)$, $u_n(t)$, and $\omega_{n,m}(t)$ still depend only weakly on t. Then, in (15.7.117) we can neglect their time dependence and as before put

$$a_m = \delta_{n,m} \ . \tag{15.7.122}$$

Then,

$$\begin{aligned}\frac{da_m}{dt} &= -\frac{2\omega}{\hbar\omega_{n,m}}(u_m, V_0u_n)\cos\omega t\, e^{-i\omega_{n,m}t} \\ &= -\frac{\omega}{\hbar\omega_{n,m}}(u_m, V_0u_n)\left[e^{-i(\omega_{n,m}-\omega)t}\right. \\ &+ \left. e^{-i(\omega_{n,m}+\omega)t}\right]\end{aligned} \tag{15.7.123}$$

or integrating

$$a_m = \frac{\omega}{i\hbar\omega_{n,m}}(u_m, V_0u_n)\left[\frac{e^{-i(\omega_{n,m}-\omega)t} - 1}{\omega_{n,m} - \omega} + \frac{e^{-i(\omega_{n,m}+\omega)t} - 1}{\omega_{n,m} + \omega}\right] \ . \tag{15.7.124}$$

Clearly for $\omega \approx \omega_{n,m}$ the adiabatic approximation breaks down since then for

$$t \leq \frac{2\pi}{\omega_{n,m} - \omega} \tag{15.7.125}$$

the quantity $T\,\partial V/\partial t$ is not small and furthermore we find that

$$\frac{e^{-i(\omega_{n,m}-\omega)t}-1}{\omega_{n,m}-\omega}\approx t \tag{15.7.126}$$

so that $|a_m|\propto t$. The same thing happens of course for $\omega\approx-\omega_{n,m}$. In fact, in either case we obtain the same result as in the perturbative treatment. Thus, we have for $\omega\approx\omega_{n,m}$

$$a_m\approx\frac{(u_m,V_0u_n)}{i\hbar}\frac{e^{-i(\omega_{n,m}-\omega)t}-1}{\omega_{n,m}-\omega} \tag{15.7.127}$$

and for $\omega\approx-\omega_{n,m}$

$$a_m\approx\frac{(u_m,V_0u_n)}{i\hbar}\frac{e^{-i(\omega_{n,m}+\omega)t}-1}{\omega_{n,m}+\omega}\ . \tag{15.7.128}$$

These are precisely the results obtained from the perturbative treatment. This should not come as a surprise since to derive this we had to assume that $V(t)$ was small. However, for the adiabatic approximation this was not required since then we only needed that

$$\left|t\,\frac{\partial V}{\partial t}\right|<<1\ . \tag{15.7.129}$$

15.8 The Sudden Approximation

As outlined in the previous section the sudden approximation is based on the fact that

$$H=H_0+\lambda V(t) \tag{15.8.130}$$

changes rapidly. In fact, it is complementary to the perturbative treatment in the following sense. The perturbative treatment is valid if the time T over which the interaction is on is relatively long compared to $1/\omega_{n,m}$. The sudden approximation is valid when the time dependence of the Hamiltonian is on for a very short time T. In general the Hamiltonian will be of the type

$$H=\begin{cases}H_0 & t<0\\ H_i & 0<t<T\\ H_1 & t>T\end{cases}\ . \tag{15.8.131}$$

This type of behaviour will occur, for example, if electrons are bound to an atom whose nucleus at $t=0$ beta-decays so that suddenly the nuclear charge is increased by one. In this case for $t<0$ the nuclear charge is Z and for $t>T$ it is $Z+1$. During the interval $0<t<T$ the Hamiltonian changes rapidly and in a very complicated manner. The advantage of the sudden approximation is that if T is sufficiently small one need not even know H_i. The approximation consists, in fact, in replacing the Hamiltonian (15.8.131) by the simpler Hamiltonian

$$H=\begin{cases}H_0 & t<0\\ H_1 & t>0\end{cases}\ . \tag{15.8.132}$$

In this case we solve the two time independent Schrödinger equations

$$H_0 u_n = E_n u_n \tag{15.8.133}$$

and

$$H_1 v_\mu = \epsilon_\mu v_\mu \ . \tag{15.8.134}$$

Then $\{u_n\}$ and $\{v_\mu\}$ both form complete (not necessarily discrete) sets. We assume that both sets are orthonormalized. Then the general solution of the time dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = H\Psi \tag{15.8.135}$$

is

$$\Psi = \sum a_n u_n e^{-iE_n t/\hbar} \quad t < 0 \tag{15.8.136}$$

$$\Psi = \sum b_\mu v_\mu e^{-i\epsilon_\mu t/\hbar} \quad t > 0 \tag{15.8.137}$$

where the sum is to be interpreted as a sum over the discrete and an integral over the continuous energy eigenvalues. Furthermore since Ψ satisfies a first order differential equation in time with a simple jump discontinuity in $\partial\Psi/\partial t$, it follows that Ψ is continuous at $t = 0$. Hence we get that

$$\sum a_n u_n = \sum b_\mu v_\mu \ . \tag{15.8.138}$$

Using the orthonormality of the $\{v_\mu\}$ we then get that

$$b_\mu = \sum a_n (v_\mu, u_n) \ . \tag{15.8.139}$$

Typically we start with the system in a pure state say u_m. Then for $t < 0$, $a_n = \delta_{n,m}$ so that

$$b_\mu = (v_\mu, u_m) \ . \tag{15.8.140}$$

Thus, starting with a pure state with energy E_m we wind up after the interaction in a superposition of states with energy ϵ_μ. Actually this is only due to our mode of description. In fact, as stated, Ψ is continuous so the state of the system has not changed and remains an eigenstate of H_0. In a sense, the Hamiltonian changes too rapidly for the system to follow and thus it remains unchanged in an eigenstate of H_0. However, for $t > 0$ it evolves according to the new Hamiltonian. This is to be contrasted with the adiabatic case where the system evolves from an eigenstate of the original Hamiltonian into an eigenstate of the final Hamiltonian.

To illustrate the application of the methods we have developped we next consider the case of a magnetic dipole in a magnetic field that oscillates or is increased either adiabatically or suddenly to some final value.

15.9 Dipole in a Time-Dependent Magnetic Field

The problem we want to consider is that of an electron in a magnetic field. We assume that all other energies of the electron such as its kinetic energy can be neglected. Since the electron has a spin **S** it has associated with it a magnetic dipole moment operator (see section 9.7)

$$\vec{\mu} = -\frac{e}{mc}\mathbf{S} \ . \tag{15.9.141}$$

Thus, with our approximations, the Hamiltonian for this electron is

$$H = -\vec{\mu}\cdot\mathbf{B} = \frac{e}{mc}\mathbf{S}\cdot\mathbf{B} = \frac{e\hbar}{2mc}\vec{\sigma}\cdot\mathbf{B} \tag{15.9.142}$$

where $\vec{\sigma}$ are the Pauli matrices (see section 9.7). We consider three separate cases in all of which

$$\mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1(t) \tag{15.9.143}$$

where $\mathbf{B}_0$ has only a z-component B_0 and $\mathbf{B}_1(t)$ has only an x-component. The three cases are:

1)

$$B_1(t) = \left\{ \begin{array}{ll} 0 & t < 0 \\ b\sin\omega t & t > 0 \end{array} \right. \ . \tag{15.9.144}$$

2)

$$B_1(t) = \left\{ \begin{array}{ll} 0 & t < 0 \\ b(1 - e^{-\alpha t})t & t > 0 \end{array} \right. \ . \tag{15.9.145}$$

3)

$$B_1(t) = \left\{ \begin{array}{ll} 0 & t < 0 \\ b & t > 0 \end{array} \right. \ . \tag{15.9.146}$$

We first solve for the steady states of the unperturbed Hamiltonian

$$H_0 = \frac{e\hbar B_0}{2mc}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{15.9.147}$$

The eigenstates are

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{15.9.148}$$

with eigenvalues $\pm\hbar\Omega_0$ respectively, where

$$\Omega_0 = \frac{eB_0}{2mc} \ . \tag{15.9.149}$$

15.9.1 Oscillatory Perturbation

In this case the Hamiltonian is

$$H = \hbar\Omega_0\sigma_3 + \hbar\Omega_1 \sin\omega t\, \sigma_1 \tag{15.9.150}$$

or

$$H = \hbar\Omega_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \hbar\Omega_1 \sin\omega t \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{15.9.151}$$

where

$$\Omega_1 = \frac{eb}{2mc} . \tag{15.9.152}$$

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = H|\Psi(t)\rangle \tag{15.9.153}$$

according to (15.3.26) with

$$E_\pm = \pm\hbar\Omega_0 \tag{15.9.154}$$

then we have only one transition frequency, namely

$$E_+ - E_- = 2\hbar\Omega_0 . \tag{15.9.155}$$

Also, we write, as in equation (15.3.26)

$$\Psi(t) = a_+(t)e^{-i\Omega_0 t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_-(t)e^{i\Omega_0 t} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{15.9.156}$$

so that (15.3.29) reduces to the two equations:

$$\frac{da_+}{dt} = -i\Omega_1 e^{2i\Omega_0 t} \sin\omega t\, a_- \tag{15.9.157}$$

$$\frac{da_-}{dt} = -i\Omega_1 e^{-2i\Omega_0 t} \sin\omega t\, a_+ . \tag{15.9.158}$$

If furthermore for $t < 0$ the electron is in the state $|+\rangle$ then the integral equations equivalent to (15.3.33) are

$$a_+ = 1 - i\Omega_1 \int_0^t e^{2i\Omega_0 t'} \sin\omega t'\, a_-(t')\, dt' \tag{15.9.159}$$

$$a_- = -i\Omega_1 \int_0^t e^{-2i\Omega_0 t'} \sin\omega t'\, a_+(t')\, dt' . \tag{15.9.160}$$

These equations are still exact. If we now iterate them to obtain successive approximations, then the zeroth order terms in Ω_1 are

$$a_+^{(0)} = 1 \qquad a_-^{(0)} = 0 \tag{15.9.161}$$

and the first order terms in Ω_1 are

$$\begin{aligned} a_+^{(1)} &= 0 \\ a_-^{(1)} &= -i\Omega_1 \int_0^t e^{-2i\Omega_0 t'} \sin\omega t'\, dt' \ . \end{aligned} \tag{15.9.162}$$

The second order terms are

$$\begin{aligned} a_+^{(2)} &= -\Omega_1^2 \int_0^t e^{2i\Omega_0 t'} \sin\omega t'\, dt' \int_0^{t'} e^{2i\Omega_0 t''} \sin\omega t''\, dt'' \\ a_-^{(2)} &= 0 \ . \end{aligned} \tag{15.9.163}$$

The first two of these equations are precisely (15.3.34) and (15.3.35) for the specific Hamiltonian (15.9.142) with $B_1(t)$ given by (15.9.144).

To first order in Ω_1 we then have

$$\begin{aligned} |a_+|^2 &= 1 \\ |a_-|^2 &= \Omega_1^2 \left| \int_0^t e^{-2i\Omega_0 t'} \sin\omega t'\, dt' \right|^2 \end{aligned} \tag{15.9.164}$$

or

$$|a_-|^2 = \frac{\Omega_1^2}{4} \left| \frac{e^{-i(2\Omega_0-\omega)t} - 1}{2\Omega_0 - \omega} - \frac{e^{-i(2\Omega_0+\omega)t} - 1}{2\Omega_0 + \omega} \right|^2 \ . \tag{15.9.165}$$

Thus, to lowest order, we find that the system will almost certainly remain in its original state $|+\rangle$ and there is only a second order term giving the probability of finding the system in the state $|-\rangle$ after a time t.

15.9.2 Slowly Varying Perturbation

Here we consider

$$B_1(t) = \begin{cases} 0 & t < 0 \\ b(1 - e^{-\alpha t}) & t \geq 0 \end{cases} \tag{15.9.166}$$

and apply the adiabatic approximation. Since the typical time constant here is given by $2\pi/\Omega_0$, this approximation will be valid if α/Ω_0 is very small. We now find

$$\frac{dB_1}{dt} = 0 \qquad t < 0 \tag{15.9.167}$$

and

$$\frac{dB_1}{dt} = \alpha b\, e^{-\alpha t} \quad t \geq 0 \ . \tag{15.9.168}$$

The Hamiltonian is

$$H = \hbar\Omega_0\sigma_3 + \hbar\Omega_1 \left(1 - e^{-\alpha t}\right) \sigma_1 \quad t \geq 0 \tag{15.9.169}$$

or

$$H = \hbar\Omega_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \hbar\Omega_1 \left(1 - e^{-\alpha t}\right) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad t \geq 0 \tag{15.9.170}$$

and

$$H = H_0 = \hbar\Omega_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \hbar\Omega_0\sigma_3 \quad t < 0 \tag{15.9.171}$$

where as before

$$\Omega_0 = \frac{eB_0}{2mc} \quad , \quad \Omega_1 = \frac{eb}{2mc} \ . \tag{15.9.172}$$

The time-dependent Schrödinger equation can be rewritten to yield the equations corresponding to (15.7.115). They are

$$\dot{a}_+ = \frac{\alpha\hbar\Omega_1}{E_+ - E_-} e^{-\alpha t} (u_+, \sigma_x u_-) \exp\left[i \int_0^t \frac{E_+ - E_-}{\hbar}\, dt'\right] a_- \tag{15.9.173}$$

$$\dot{a}_- = \frac{\alpha\hbar\Omega_1}{E_- - E_+} e^{-\alpha t} (u_-, \sigma_x u_+) \exp\left[i \int_0^t \frac{E_- - E_+}{\hbar}\, dt'\right] a_+ \tag{15.9.174}$$

where E_+, E_- and u_+, u_- are solutions of the Hamiltonian for $t \geq 0$

$$H(t)u_\pm(t) = E_\pm(t)u_\pm(t) \ . \tag{15.9.175}$$

These equations are still exact and also far too complicated to solve. The adiabatic approximation consists in choosing the $a_\pm$, $u_\pm$ on the right hand side of (15.9.173) and (15.9.174) as well as the energies $E_\pm$ as constants. Thus, if for example the initial state at $t = 0$ was u_+ then on the right side of (15.9.174) and (15.9.173) we have $a_+ = 1$, $a_- = 0$. This yields,

$$\begin{aligned} \frac{da_+}{dt} &= 0 \\ \frac{da_-}{dt} &= -\frac{\alpha\hbar\Omega_1}{2\hbar\Omega_0} e^{-\alpha t} e^{-2i\Omega_0 t} \end{aligned} \tag{15.9.176}$$

or, integrating

$$\begin{aligned} a_+ &= 1 \\ a_- &= \frac{\alpha}{2(\alpha + 2i\Omega_0)} \frac{\Omega_1}{\Omega_0} \left[e^{-(\alpha+2i\Omega_0)t} - 1\right] \\ &\approx -i\frac{\alpha\Omega_1}{4\Omega_0^2} \left[e^{-(\alpha+2i\Omega_0)t} - 1\right] \ . \end{aligned} \tag{15.9.177}$$

So, to lowest order in α/Ω_0, the state at time t is given by

$$\Psi(t) = \begin{pmatrix} 1 \\ -i\frac{\alpha\Omega_1}{4\Omega_0^2} \left[e^{-(\alpha+2i\Omega_0)t} - 1\right] \end{pmatrix} \ . \tag{15.9.178}$$

This shows that the state evolves slowly from its original state u_+ to a mixture of u_+ and u_-.

15.9.3 Sudden Approximation

In this case the Hamiltonian is

$$H = \hbar\Omega_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = H_0 = \hbar\Omega_0\sigma_3 \quad t < 0 \tag{15.9.179}$$

and

$$H = \hbar\Omega_0\sigma_3 + \hbar\Omega_1\sigma_1 = H_1 \quad t > 0 \tag{15.9.180}$$

or

$$H = \hbar\Omega_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \hbar\Omega_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = H_1 \quad t > 0\,. \tag{15.9.181}$$

The eigenstates of H_0 are, as before,

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{15.9.182}$$

with

$$E_\pm = \pm\hbar\Omega_0\ . \tag{15.9.183}$$

If we call the eigenfunctions and eigenvalues of H_1, v_μ and ϵ_μ respectively, we have for

$$v_\mu = \begin{pmatrix} a \\ b \end{pmatrix} \tag{15.9.184}$$

the eigenvalue equation

$$\hbar \begin{pmatrix} \Omega_0 & \Omega_1 \\ \Omega_1 & -\Omega_0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \epsilon_\mu \begin{pmatrix} a \\ b \end{pmatrix}\ . \tag{15.9.185}$$

The eigenvalues are:

$$\epsilon_\pm = \pm\hbar\sqrt{\Omega_0^2 + \Omega_1^2}\,. \tag{15.9.186}$$

Calling

$$\Omega^2 = \Omega_0^2 + \Omega_1^2 \tag{15.9.187}$$

we have

$$\epsilon_\pm = \pm\hbar\Omega\ . \tag{15.9.188}$$

We then get:

$$v_\pm = \frac{1}{\sqrt{2\Omega(\Omega - \Omega_0)}} \begin{pmatrix} \Omega_1 \\ \pm\Omega - \Omega_0 \end{pmatrix}\ . \tag{15.9.189}$$

If at $t = 0$ the system is in the state $u_+ = |+\rangle$, then for $t > 0$ the system will be in the state

$$\Psi(t) = b_+ v_+ e^{-i\Omega t} + b_- v_- e^{i\Omega t} \tag{15.9.190}$$

where $b_\pm$ are determined by the initial condition

$$u_+ = b_+ v_+ + b_- v_- \tag{15.9.191}$$

or

$$b_\pm = (v_\pm, u_+) . \tag{15.9.192}$$

Thus,

$$b_\pm = \frac{\Omega_1}{\sqrt{2\Omega(\Omega \mp \Omega_0)}} . \tag{15.9.193}$$

Combining these results we get

$$\Psi(t) = \begin{pmatrix} \cos \Omega t - i\frac{\Omega_0}{\Omega} \sin \Omega t \\ -i\frac{\Omega_0}{\Omega} \sin \Omega t \end{pmatrix} . \tag{15.9.194}$$

This wavefunction Ψ is an *exact* solution of the Hamiltonian (15.9.179) and (15.9.180) with the initial condition $\Psi(0) = u_+$. The approximation is made in writing the Hamiltonian (15.9.179), (15.9.180) in the first place.

15.10 Two-Level Systems

In section 13.2 we considered a two-dimensional degenerate subspace. We also considered such a two-dimensional system in the previous sections. We now re-examine such two-dimensional systems in general. Two-dimensional systems frequently provide a good approximation to a real physical system other than a spin 1/2 system as discussed above. An important example is provided by an ionized hydrogen molecule where the electron may be found in the ground state around either of the two protons.

To bring out the essential features of such a system we write the Hamiltonian in the following form:

$$H = H_0 + V \tag{15.10.195}$$

where

$$H_0 = \begin{pmatrix} E - \epsilon & 0 \\ 0 & E + \epsilon \end{pmatrix} \quad , \quad \epsilon > 0 \tag{15.10.196}$$

with eigenvalues $E_1 = E - \epsilon$, $E_2 = E + \epsilon$ and eigenvectors

$$\chi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad \chi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{15.10.197}$$

The interaction term V is written

$$V = \begin{pmatrix} 0 & H_{12} \\ H_{21} & 0 \end{pmatrix} \quad , \quad H_{12} = H_{21}^* = |v|e^{i\alpha} . \tag{15.10.198}$$

We have written V without any diagonal elements since these could always be absorbed in a redefinition of H_0. The eigenvalues of the full Hamiltonian H are:

$$E_\pm = E \pm \sqrt{\epsilon^2 + |v|^2} \ . \tag{15.10.199}$$

We now define, as in section 13.2,

$$\tan\theta = \frac{|v|}{\epsilon} \ . \tag{15.10.200}$$

The eigenvectors of H corresponding to $E_\pm$ are (up to an arbitrary choice of phase), given by:

$$\psi_- = \begin{pmatrix} \cos\theta/2\, e^{i\alpha} \\ -\sin\theta/2 \end{pmatrix} \ , \quad \psi_+ = \begin{pmatrix} \sin\theta/2\, e^{i\alpha} \\ \cos\theta/2 \end{pmatrix} \ . \tag{15.10.201}$$

An interesting point to note here is that the eigenvectors do not depend on the "average unperturbed energy" E but only on the difference ϵ. Furthermore, we see that

$$|E_+ - E_-| > |E_2 - E_1| = 2\epsilon \ . \tag{15.10.202}$$

Thus, the effect of the interaction is to "repel" the two energy levels. This result was also found in section 13.2.

The time evolution of this system is, as usual, governed by the Schrödinger equation.

$$i\hbar\frac{\partial\Psi}{\partial t} = H\Psi \tag{15.10.203}$$

so that

$$\Psi(t) = e^{-iHt/\hbar}\Psi(0) \ . \tag{15.10.204}$$

Now the most general solution of equation (15.10.203) is given by:

$$\Psi(t) = c_1(t)\chi_1 + c_2(t)\chi_2 \ . \tag{15.10.205}$$

If we use the spectral decomposition (see problem 6.5) of H in the form

$$\begin{aligned} H &= E_-\psi_-\psi_-^\dagger + E_+\psi_+\psi_+^\dagger \\ &= E_-\begin{pmatrix} \cos^2\theta/2 & -\sin\theta/2\,\cos\theta/2\, e^{i\alpha} \\ -\sin\theta/2\,\cos\theta/2\, e^{-i\alpha} & \sin^2\theta/2 \end{pmatrix} \\ &+ E_+\begin{pmatrix} \sin^2\theta/2 & \sin\theta/2\,\cos\theta/2\, e^{i\alpha} \\ \sin\theta/2\,\cos\theta/2\, e^{-i\alpha} & \cos^2\theta/2 \end{pmatrix} \end{aligned} \tag{15.10.206}$$

then we also have

$$e^{-iHt/\hbar} = e^{-iE_-t/\hbar}\psi_-\psi_-^\dagger + e^{-iE_+t/\hbar}\psi_+\psi_+^\dagger \ . \tag{15.10.207}$$

So we can immediately write

$$\begin{aligned} \Psi(t) &= e^{-iHt/\hbar}\Psi(0) \\ &= e^{-iE_-t/\hbar}\psi_-(\psi_-, \Psi(0)) + e^{-iE_+t/\hbar}\psi_+(\psi_+, \Psi(0)) \ . \end{aligned} \tag{15.10.208}$$

Using this equation we are ready to compute a particularly important result which is sometimes known as the *Rabi Formula.* What we are interested in is the probability P_{12} that, the system which starts out at time $t = 0$ in the state

$$\Psi(0) = \chi_1 \tag{15.10.209}$$

is at time t in the state χ_2. We don't need to compute the probabilty P_{11} that, the system which starts out at time $t = 0$ in the state χ_1 is again in the same state χ_1 at time t since this is given by $1 - P_{12}$. We, theore, only need to compute

$$P_{12} = \left|\left(\chi_2, e^{-iHt/\hbar}\chi_1\right)\right|^2 . \tag{15.10.210}$$

To accomplish this we first express χ_1 in terms of ψ_+, ψ_- and use (15.10.207). Now, from (15.10.201) we get

$$\begin{aligned} \chi_1 &= e^{-i\alpha}(\cos\theta/2\,\psi_- + \sin\theta/2\,\psi_+) \\ \chi_2 &= -\sin\theta/2\,\psi_- + \cos\theta/2\,\psi_+ . \end{aligned} \tag{15.10.211}$$

Theore,

$$e^{-iHt/\hbar}\chi_1 = e^{-i\alpha}\left(\cos\theta/2\,e^{-iE_-t/\hbar}\,\psi_- + \sin\theta/2\,e^{-iE_+t/\hbar}\,\psi_+\right) . \tag{15.10.212}$$

Hence we find,

$$\begin{aligned} P_{12} &= \left|\cos\theta/2\,e^{-iE_-t/\hbar}(\chi_2,\psi_-) + \sin\theta/2\,,e^{-iE_+t/\hbar}(\chi_2,\psi_+)\right|^2 \\ &= \frac{1}{4}\sin^2\theta\left|e^{-iE_+t/\hbar} - e^{-iE_-t/\hbar}\right|^2 \\ &= \frac{1}{2}\sin^2\theta\left[1 - \cos\left(\frac{E_+ - E_-}{\hbar}t\right)\right] . \end{aligned} \tag{15.10.213}$$

So finally

$$P_{12} = \sin^2\theta\,\sin^2\left(\frac{E_+ - E_-}{2\hbar}t\right) . \tag{15.10.214}$$

Substituting for $E_+ - E_-$ and $\sin^2\theta$ from (15.10.199) and (15.10.200) we get

$$P_{12} = \frac{|v|^2}{\epsilon^2 + |v|^2}\sin^2\left(\frac{\sqrt{\epsilon^2 + |v|^2}}{\hbar}t\right) . \tag{15.10.215}$$

This is the *Rabi Formula.* It shows that as the system evolves in time, it oscillates between the levels χ_1 and χ_2 with angular frequency

$$\omega = \frac{\sqrt{\epsilon^2 + |v|^2}}{\hbar} . \tag{15.10.216}$$

An example of such a two-level system, as stated at the beginning of this section, is given by the ground state of an H_2^+ molecule. There the electron oscillates between the two states in which it is localized first around one proton and then around the other. In this case, the unperturbed energy levels are degenerate and $\epsilon = 0$.

15.11 Berry's Phase

In section 15.7 we studied the adiabatic approximation and assumed that we could ignore a certain phase factor $e^{i\gamma(t)}$ so that we could set

$$\left(u_m, \frac{\partial u_m}{\partial t}\right) = 0 \ . \tag{15.11.217}$$

This was equation (15.7.114). Of course, the dynamical phase factor

$$\exp\left(-\frac{i}{\hbar}\int_0^t E_n(t')\,dt'\right)$$

was, as always, present. In 1984, M.V. Berry [15.3] showed that if the adiabatically varying Hamiltonian $H(t)$ is periodic such that

$$H(t_1) = H(t_2) \tag{15.11.218}$$

then, for a time interval from t_1 to t_2 one obtains, in addition to the dynamical phase above, an additional *observable* topological phase factor $e^{i\gamma_n}$. We now present his argument.

Consider the Hamiltonian $H(t)$ and assume that the time dependence of $H(t)$ is through a set of k slowly varying parameters

$$\underline{R}(t) = (R_1(t), R_2(t), \cdots, R_k(t)) \tag{15.11.219}$$

that are periodic so that for times t_1 , t_2

$$\underline{R}(t_2) = \underline{R}(t_1) \tag{15.11.220}$$

and

$$\left|\frac{1}{R_i(t)}\frac{dR_i(t)}{dt}\right| \ll \left|\frac{1}{E_n(t)}\frac{dE_n(t)}{dt}\right| \ . \tag{15.11.221}$$

This last inequality is just a more specific statement of the adiabatic condition.

The Schrödinger equation now reads

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = H(\underline{R}(t))|\Psi(t)\rangle \ . \tag{15.11.222}$$

We again have instantaneous eigenkets, $|n, \underline{R}(t)\rangle$, of $H(\underline{R}(t))$ with instantaneous energy eigenvalues $E_n(\underline{R}(t))$.

$$H(\underline{R}(t))|n, \underline{R}(t)\rangle = E_n(\underline{R}(t))|n, \underline{R}(t)\rangle \ . \tag{15.11.223}$$

A solution of the time-dependent Schrödinger equation can therefore be written

$$|\Psi(t)\rangle = \exp\left(-\frac{i}{\hbar}\int_0^t E_n(t')\,dt'\right) e^{i\gamma_n(t)}\,|n, \underline{R}(t)\rangle \ . \tag{15.11.224}$$

Here, in addition to the usual dynamical first exponential factor we have included the explicit phase factor $e^{i\gamma_n(t)}$ which was argued away in section 15.7.

Berry's crucial discovery was that, for certain situations, $\gamma_n(t)$ is non-integrable, that is, $\gamma_n(t)$ can not be written as a function of $\underline{R}$ and, in particular, is not single-valued under completion of a period from t_1 to t_2. Thus,

$$\gamma_n(t_1) \neq \gamma_n(t_2) \ . \tag{15.11.225}$$

As in section 15.7, the explicit expression for $\gamma_n(t)$ is obtained by substituting the expression (15.11.224) into the Schrödinger equation. This yields

$$\frac{d\gamma_n}{dt} = i\langle n, \underline{R}(t)|n, \dot{\underline{R}}(t)\rangle \ . \tag{15.11.226}$$

The term $|n, \dot{\underline{R}}(t)\rangle$ may be rewritten

$$|n, \dot{\underline{R}}(t)\rangle = \sum_i^k \frac{\partial}{\partial R_i}|n, \underline{R}(t)\rangle \dot{R}_i = (\nabla_R|n, \underline{R}(t)\rangle) \cdot \dot{\underline{R}} \ . \tag{15.11.227}$$

Here we have introduced the symbol ∇_R as an obvious generalization of the gradient to k dimensions. Using this expression we find that the net phase change of $|\Psi(t)\rangle$ in one period (say from t_1 to t_2) is given by

$$|\Psi(t_2)\rangle = e^{i\gamma_n(C)} \exp\left(-\frac{i}{\hbar}\int_{t_1}^{t_2} E_n(t')\,dt'\right)|\Psi(t_1)\rangle \tag{15.11.228}$$

where the *geometrical phase change* is

$$\begin{aligned} \gamma_n(C) &= i\int_{t_1}^{t_2} \langle n, \underline{R}|\nabla_R|n, \underline{R}\rangle \cdot \frac{d\underline{R}}{dt}\,dt \\ &= i\oint_C \langle n, \underline{R}|\nabla_R|n, \underline{R}\rangle \cdot d\underline{R} \,. \end{aligned} \tag{15.11.229}$$

This shows explicitly that $\gamma_n(C)$ is given by integrating around a closed loop in the parameter space (i.e. from $\underline{R}(t_1)$ to $\underline{R}(t_2) = \underline{R}(t_1)$). We now go one step further and use Stoke's theorem to convert the line integral into a surface integral in parameter space. The surface integral runs over *any* surface having the closed curve C, in parameter space, as a boundary.

At this stage we also restrict the argument to the more intuitive and physically most interesting case of a three-dimensional parameter space. With this restriction it is useful to introduce a vector field $\mathbf{A}(\underline{R})$ by writing

$$i\langle n, \underline{R}|\nabla_R|n, \underline{R}\rangle \equiv \mathbf{A}(\mathbf{R}) \tag{15.11.230}$$

where we are now also writing $\mathbf{R}$ instead of $\underline{R}$. We then have the more suggestive formula

$$\begin{aligned} \gamma_n(C) &= \oint_C \mathbf{A}(\mathbf{R}) \cdot d\mathbf{R} \\ &= \int_S (\nabla \times \mathbf{A}) \cdot d\mathbf{S} \,. \end{aligned} \tag{15.11.231}$$

This use of a "vector potential" **A** shows that if we redefine the phase of $|n,\underline{R}\rangle$ so that we introduce

$$|n',\underline{R}\rangle = e^{i\Lambda(\mathbf{R})}\,|n,\underline{R}\rangle \tag{15.11.232}$$

as we did earlier to get rid of the term $\langle n,\underline{R}(t)|n,\dot{\underline{R}}(t)\rangle$, the net effect is merely to produce a change analogous to a gauge transformation on **A**, namely

$$\mathbf{A}\to\mathbf{A}'=\mathbf{A}-\nabla\Lambda\ . \tag{15.11.233}$$

Then, as (15.11.231) shows, the geometric phase $\gamma_n(C)$ is left unchanged since

$$\int_S \nabla\times(\nabla\Lambda)\cdot d\mathbf{S}=\oint_C \nabla\Lambda\cdot d\mathbf{R}=\oint_C d\Lambda=0\ . \tag{15.11.234}$$

Thus, in this case it is impossible to get rid of this phase by a redefinition as we did before. Equation (15.11.231) furthermore also shows that the geometric phase $\gamma_n(C)$ is just the flux through the surface S of a vector field $\mathbf{V}_n$ given by the curl of **A**.

$$\mathbf{V}_n=\nabla\times\mathbf{A} \tag{15.11.235}$$

and

$$\gamma_n(C)=\int_S \mathbf{V}_n\cdot d\mathbf{S}\ . \tag{15.11.236}$$

To proceed we first rewrite the *real* vector field $\mathbf{V}_n$. Here we use the obvious notation that

$$\nabla_R|n,\mathbf{R}\rangle=|\nabla_R(n,\mathbf{R})\rangle \tag{15.11.237}$$

$$\nabla_R\langle n,\mathbf{R}|=\langle\nabla_R(n,\mathbf{R})|\ . \tag{15.11.238}$$

Then,

$$\begin{aligned}\mathbf{V}_n &= i\nabla_R\times\langle n,\mathbf{R}|\nabla_R|n,\mathbf{R}\rangle\\ &= i\langle\nabla_R(n,\mathbf{R})|\times\nabla_R(n,\mathbf{R})\rangle\\ &= -\Im\langle\nabla_R(n,\mathbf{R})|\times\nabla_R(n,\mathbf{R})\rangle\\ &= -\Im\sum_m\langle\nabla_R(n,\nabla_R)|m,\mathbf{R}\rangle|\times\langle m,\mathbf{R}|\nabla_R(n,\mathbf{R})\rangle\ .\end{aligned} \tag{15.11.239}$$

Next, we use the instantaneous Schrödinger equation (15.11.223) and operate on it with ∇_R. Then, we take matrix elements with $\langle m,\mathbf{R}|\quad m\neq n$ to get:

$$\begin{aligned}&\langle m,\mathbf{R}|\nabla_R H(\mathbf{R})|n,\mathbf{R}\rangle+E_m(\mathbf{R})\langle m,\mathbf{R}|\nabla_R(n,\mathbf{R})\rangle\\ =\ &E_n(\mathbf{R})\langle m,\mathbf{R}|\nabla_R(n,\mathbf{R})\rangle\ .\end{aligned} \tag{15.11.240}$$

Thus, we find

$$\langle m,\mathbf{R}|\nabla_R(n,\mathbf{R})\rangle=\frac{\langle m,\mathbf{R}|\nabla_R H(\mathbf{R})|n,\mathbf{R}\rangle}{E_n(\mathbf{R})-E_m(\mathbf{R})}\qquad m\neq n\ . \tag{15.11.241}$$

Furthermore, we have

$$\langle n,\mathbf{R}|n,\mathbf{R}\rangle = 1 \tag{15.11.242}$$

so that

$$\langle n,\mathbf{R}|\nabla_R(n,\mathbf{R})\rangle + \langle\nabla_R(n,\mathbf{R})|n,\mathbf{R}\rangle = 0 \ . \tag{15.11.243}$$

This shows that

$$\Re\langle n,\mathbf{R}|\nabla_R(n,\mathbf{R})\rangle = 0 \tag{15.11.244}$$

and means that $\langle n,\mathbf{R}|\nabla_R(n,\mathbf{R})\rangle$ is purely imaginary. Theore, in the sum over intermediate states in (15.11.239) the term with $m = n$ (which is purely real) does not contribute. Hence, we finally get

$$\mathbf{V}_n(\mathbf{R}) = -\Im\sum_{m\neq n}\frac{\langle n,\mathbf{R}|\nabla_R H(\mathbf{R})|m,\mathbf{R}\rangle\times\langle m,\mathbf{R}|\nabla_R H(\mathbf{R})|n,\mathbf{R}\rangle}{(E_n(\mathbf{R})-E_m(\mathbf{R}))^2} \ . \tag{15.11.245}$$

For the loop integral

$$\oint_C \mathbf{V}_n\cdot d\mathbf{R}$$

to give a non-zero result requires that $\mathbf{V}_n$ have a singularity inside the loop. In other words, $\mathbf{V}_n$ must have a source. Clearly this will occur if we find that for some value of t, the parameter $\mathbf{R}(t)$ is such that degeneracy of the instantaneous energy eigenvalues occurs so that $E_n(\mathbf{R}(t)) = E_m(\mathbf{R}(t))$. This means, that in this case, the occurence of a non-zero Berry's phase is connected with the occurence of degeneracy. Berry's phase is then the flux associated with the source of the field $\mathbf{V}_n$.

A simple specific example of this kind is demonstrated in problems 15.10 and 15.11 below by the Hamiltonian of a spin 1/2 particle with a magnetic moment μ in the field of a time-dependent "magnetic monopole".

15.12 Problems

15.1 A particle is in the ground state of the Hamiltonian

$$H = \frac{p^2}{2m} + V$$

where

$$V = \begin{cases} 0 & x < -a \ , \ x > a \\ -V_0 & -a < x < a \end{cases} \ .$$

Find the transition probability per unit time to a state of energy $E_k > 0$, due to a perturbation

$$H'(t) = v\, e^{-x^2/\alpha^2}\sin\omega t$$

where v is a constant and $\alpha << a$. You may use the result of problem 8.1.

15.2 The deuteron is an s-wave ($l = 0$) bound state of a proton and neutron with a binding energy of 2.226 MeV. It is well approximated as a bound state in a square well of depth $V_0 = 36.2$ MeV and a width $a = 2.02 \times 10^{-13}$ cm. Using these data, compute the probability for photo-disintegration of the deuteron. Assume the incident photon can be approximated by a perturbation

$$V = \begin{cases} e\mathbf{A} \cdot \mathbf{r} \sin \omega t & t > 0 \\ 0 & t < 0 \end{cases}$$

where **A** is a constant vector of magnitude about 1×10^3 V/cm. Use whatever other approximations seem reasonable.

15.3 An atom is initially in the ground state of a simple harmonic oscillator

$$H = \hbar\omega\, a^\dagger a \ .$$

At $t = 0$ a perturbation

$$V' = \hbar\Omega(a^\dagger + a)$$

is turned on. Find the transition probability to any exd state of the system for $t > 0$. What is the probability that the atom remains in its ground state for $t > 0$?

15.4 Repeat problem 15.3 with

$$V' = \hbar\Omega\, a^\dagger a \ .$$

15.5 An atom has two energy levels $\pm\hbar\Omega$. A weak disturbance $V(t)$ connecting these two levels and varying periodically in time such that

$$\langle 1|V(t)|2\rangle = \hbar\Omega_1 \sin\omega t$$

is turned on at $t = 0$.

a) Find a model Hamiltonian for this system.

b) If the atom was originally in its ground state, estimate the probability $P(t)$ that it is in its exd state at time t.

15.6 An hydrogen atom in an exd state $|n, l, m\rangle$ is perturbed by a uniform electric field. If the interaction can be written

$$V(t) = \begin{cases} e\mathbf{E} \cdot \mathbf{r}\, 2 \sin\omega t & 0 < t < T \\ 0 & t < 0 \ , \ t > T \end{cases} .$$

Find an expression for the transition probabilities to a definite lower level. Do not attempt to evaluate the radial integrals. This is how intensities of spectral lines can be computed. You may use the results of problem 9.4. See also reference [15.1].

15.7 A particle is in the ground state in a one dimensional box of length L. Suddenly, at $t = 0$ the box expands (symmetrically) to a length $2L$ leaving the wavefunction undisturbed. Calculate the probability that at some later time t the particle will be found in the ground state of the expanded box.

15.8 A particle is bound by a simple harmonic oscillator potential and is in the first exd state. If a perturbation

$$V' = \lambda x$$

is turned on at time $t = 0$, find the probability that the particle will be in the new ground state for times $t > 0$.

15.9 An atom has two energy levels of energy $\pm E$. So the Hamiltonian may be written

$$H = E\sigma_3$$

where

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

If this atom is in the ground state at time $t = 0$ and a perturbation

$$H' = V \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

is turned on find the probability that this atom is still in its ground state at some later time t.

15.10 Consider the Hamiltonian

$$H(\mathbf{R}(t)) = -\frac{\mu}{2}\sigma \cdot \mathbf{R}(t)$$

where

$$\mathbf{B}(t) = B_0 \left[\sin\theta \cos\omega t \, \mathbf{e}_x + \sin\theta \sin\omega t \, \mathbf{e}_y + \cos\theta \, \mathbf{e}_z\right]$$

so that

$$\mathbf{B}(t + 2\pi/\omega) = \mathbf{B}(t) .$$

Show that if $\omega << \mu B_0$ then

$$\langle \Psi(t = 2\pi/\omega) | \Psi(t = 0) \rangle = \exp\left(i\frac{2\pi}{\omega}\mu B_0\right) \exp[-i\pi(1 - \cos\theta)] .$$

Hint: $\frac{2\pi}{\omega}\mu B_0$ is just the dynamical phase $-\int_0^{2\pi/\omega} E(t)\, dt$ and

$$-\pi(1 - \cos\theta) = -\Delta\Omega/2$$

is the geometric phase (Berry's phase).

15.11 Consider a neutron (charge = 0 , magnetic moment $\mu = -g\frac{e\hbar}{2m}\sigma$ in a magnetic field consisting of a uniform component B in the z-direction and a component $b(t)$ rotating in the $x-y$ plane. Assume that at $t=0$ the neutron has its spin pointing in the negative z-direction. Find $P_+(t)$, the probability that at time $t>0$ the neutron has spin $\hbar/2$ in the positive z-direction. Find the condition between B and b such that the amplitude of oscillation of the spin between $+\hbar/2$ and $-\hbar/2$ is a maximum.

Bibliography

[15.1] The computation of transition probabilities for atomic systems is carried out in
E.U. Condon and G.H. Shortley, *The Theory of Atomic Spectra* - Cambridge University Press (1963).

[15.2] A very concise, but very readable treatment of time-dependent perturbation theory is to be found in sections 44-46 of Dirac's book (reference [11.2]).

[15.3] M.V. Berry, Proc Roy Soc Lond, **A392**, 45-57 (1984).

Chapter 16

Particle in a Uniform Magnetic Field

16.1 Introduction

In this chapter we apply the techniques, developed in the previous chapters, to some specific, interesting problems. Interactions with the electromagnetic field are of particular interest to us. Since these interactions are written in terms of electromagnetic potentials we study the consequences of choosing different potentials (gauges) that give rise to the same electromagnetic field. We go on to include interactions between the electromagnetic field and the magnetic moment of an electron due to its spin as well as its orbital angular momentum.

In particular, we also study the splitting of spectral lines due to the magnetic field - the Zeeman effect. To carry out computations for this effect requires the study of the addition of angular momenta. This problem constitutes a major portion of this chapter.

16.2 Gauge Transformations

In Chapter 3, equation (3.8.80), we wrote the Hamiltonian H for a charged particle in an electromagnetic field. There we found that for a static field described by a scalar potential ϕ and a vector potential $\mathbf{A}$ such that

$$\mathbf{E} = -\nabla\phi \tag{16.2.1}$$

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{16.2.2}$$

the Hamiltonian for a particle of mass m and charge q is

$$H = \frac{1}{2m}\left(\mathbf{p} - q/c\,\mathbf{A}\right)^2 + q\phi\ . \tag{16.2.3}$$

Of course ϕ and $\mathbf{A}$ do not describe the electromagnetic field $(\mathbf{E}, \mathbf{B})$ uniquely since both $\mathbf{E}$ and $\mathbf{B}$ are left unchanged by gauge transformations. An example of such a transformation is

$$\begin{aligned} \mathbf{A} &\to \mathbf{A}' = \mathbf{A} + \nabla\Lambda \\ \phi &\to \phi' = \phi \end{aligned} \tag{16.2.4}$$

where Λ is an arbitrary time-independent scalar field. Clearly

$$\nabla \times \mathbf{A}' = \nabla \times \mathbf{A} \tag{16.2.5}$$

since the curl of a gradient vanishes. Thus, both $\mathbf{A}$ and $\mathbf{A}'$ describe the same magnetic field $\mathbf{B}$.

If we now consider the Schrödinger equation with the Hamiltonian (16.2.3) then, to keep this equation invariant under the gauge transformation (16.2.4) so that the physics does not depend on our choice of gauge, the phase of the wave-function ψ must change

$$\psi \to \psi' = e^{i\chi}\,\psi \ . \tag{16.2.6}$$

This is the active view of gauge transformations. With a proper choice of χ, the pair of transformations (16.2.4), (16.2.6) leave the Schrödinger equation

$$\frac{1}{2m}\left(\mathbf{p} - q/c\,\mathbf{A}\right)^2 \psi + q\phi\psi = E\psi \tag{16.2.7}$$

unchanged. To see this we replace $\mathbf{A}$ and ψ by $\mathbf{A}'$ and ψ' respectively to get

$$\frac{1}{2m}\left(\mathbf{p} - q/c\,\mathbf{A}' + q/c\,\nabla\Lambda\right)^2 e^{-i\chi}\,\psi' + q\phi\, e^{-i\chi}\,\psi' = E\, e^{-i\chi}\,\psi' \ . \tag{16.2.8}$$

Now writing out the action of the operator $\mathbf{p}$ we find

$$\left(\mathbf{p} - q/c\,\mathbf{A}' + q/c\,\nabla\Lambda\right) e^{-i\chi}\,\psi' = e^{-i\chi}\left[-\hbar\nabla\chi + \mathbf{p} - \frac{q}{c}\left(\mathbf{A} - \nabla\Lambda\right)\right]\psi' . \tag{16.2.9}$$

Thus, if we choose

$$\chi = \frac{q}{\hbar c}\Lambda \tag{16.2.10}$$

then equation (16.2.9) shows that

$$\left(\mathbf{p} - q/c\,\mathbf{A}' + q/c\,\nabla\Lambda\right)\, e^{-i(q/\hbar c)\Lambda}\,\psi' = e^{-i(q/\hbar c)\Lambda}\left[\mathbf{p} - \frac{q}{c}\mathbf{A}'\right]\psi' . \tag{16.2.11}$$

Operating on (16.2.11) once more with $(\mathbf{p} - q/c\,\mathbf{A}' + q/c\,\nabla\Lambda)$ we find that for χ given by (16.2.10) the Schrödinger equation (16.2.8) reduces to

$$e^{-i(q/\hbar c)\Lambda}\left[\frac{1}{2m}\left(\mathbf{p} - q/c\,\mathbf{A}'\right)^2 \psi' + q\phi'\psi'\right] = e^{-i(q/\hbar c)\Lambda}\, E\,\psi' . \tag{16.2.12}$$

After cancelling the phase factor $e^{-i(q/\hbar c)\Lambda}$ we recognize this as the original Schrödinger equation (16.2.7) with primes on all the quantities. Thus, we have found that under the local gauge transformation

$$\begin{aligned} \mathbf{A} &\rightarrow \mathbf{A}' = \mathbf{A} + \nabla\Lambda \\ \phi &\rightarrow \phi' = \phi \end{aligned} \tag{16.2.13}$$

$$\psi \rightarrow \psi' = e^{(iq/\hbar c)\Lambda}\,\psi \tag{16.2.14}$$

the Schrödinger equation

$$\frac{1}{2m}\left(\mathbf{p} - q/c\,\mathbf{A}\right)^2 \psi + q\phi\psi = E\psi \tag{16.2.15}$$

remains unchanged in form. Furthermore, the charge and current densities

$$\rho = \psi^*\psi \tag{16.2.16}$$

and

$$\mathbf{j} = \frac{\hbar}{2im}\left[\psi^*\left(\nabla + \frac{iq}{\hbar}\mathbf{A}\right)\psi - \psi\left(\nabla - \frac{iq}{\hbar}\mathbf{A}\right)\psi^*\right] \tag{16.2.17}$$

also remain unchanged. This is a very important physical result since it shows that whatever set of potentials $(\phi\,,\,\mathbf{A})$ we choose, the resultant Schrödinger equation does not depend on our choice as long as these different potentials are connected by local gauge transformations or, what is the same thing, describe the same electromagnetic field $(\mathbf{E}\,,\,\mathbf{B})$. The adjective "local" in the above discussion simply ers to the fact that the gauge field Λ depends on the coordinates x in a local manner. If the electromagnetic field $(\mathbf{E}\,,\,\mathbf{B})$ is time-dependent then the potentials $(\phi\,,\,\mathbf{A})$ are time dependent and

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t} \tag{16.2.18}$$

$$\mathbf{B} = \nabla \times \mathbf{A}\ . \tag{16.2.19}$$

In this case we must consider the invariance of the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = \frac{1}{2m}\left(\mathbf{p} - q/c\,\mathbf{A}\right)^2\Psi + q\phi\Psi \tag{16.2.20}$$

under time-dependent local gauge transformations. A computation similar to equations (16.2.9) and (16.2.11) (see problem 16.4) shows that the set of transformations

$$\begin{aligned} \mathbf{A} &\rightarrow \mathbf{A}' = \mathbf{A} + \nabla\Lambda \\ \phi &\rightarrow \phi' = \phi - \frac{1}{c}\frac{\partial\Lambda}{\partial t} \\ \Psi &\rightarrow \Psi' = e^{i(q/\hbar c)\Lambda}\,\Psi \end{aligned} \tag{16.2.21}$$

leaves the time-dependent Schrödinger equation (16.6) unchanged or invariant. In this discussion Λ is an arbitrary time-dependent scalar field.

Local gauge invariance, as discussed above, has not only the important physical consequence that the physics of interactions with the electromagnetic fields does not depend on the choice of gauge, but has proved to be an important guiding principle in modern theories of elementary particles.

As an example, that we use in the next section, we consider the vector potential for a constant magnetic field **B**. Using the vector identity

$$\nabla \times (\mathbf{F} \times \mathbf{G}) = \mathbf{F}(\nabla \cdot \mathbf{G}) - \mathbf{G}(\nabla \cdot \mathbf{F}) - (\mathbf{F} \cdot \nabla)\mathbf{G} + (\mathbf{G} \cdot \nabla)\mathbf{F} \qquad (16.2.22)$$

we see that if we choose

$$\mathbf{A} = -\frac{1}{2}\mathbf{r} \times \mathbf{B} \qquad (16.2.23)$$

then

$$\nabla \times \mathbf{A} = -\frac{1}{2}\left[\mathbf{r}(\nabla \cdot \mathbf{B}) - \mathbf{B}(\nabla \cdot \mathbf{r}) - (\mathbf{r} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{r}\right] = \mathbf{B}\ . \qquad (16.2.24)$$

Thus, a possible choice of vector potentials is given by (16.2.23). This is sometimes called the *symmetric gauge.*

For convenience we now choose our coordinate system so that the z-axis is parallel to **B**. Then,

$$\mathbf{B} = (0, 0, B) \qquad (16.2.25)$$

and

$$\mathbf{A} = -\frac{B}{2}(y, -x, 0)\ . \qquad (16.2.26)$$

If we now perform a gauge transformation of the type given by equation (16.2.4) with the gauge function

$$\Lambda = -\frac{B}{2}\,xy \qquad (16.2.27)$$

then

$$\nabla\Lambda = -\frac{B}{2}(y, x, 0) \qquad (16.2.28)$$

and

$$\mathbf{A}' = \mathbf{A} + \nabla\Lambda = -B(y, 0, 0)\ . \qquad (16.2.29)$$

The potential $\mathbf{A}'$ is just as good as the potential **A** for computing the magnetic field **B**. The choice of potential $\mathbf{A}'$ is called the *Landau gauge* [16.1]. In the next section we consider the motion of an electron in a uniform magnetic field and use both forms of the vector potential to see that they yield the same result.

For computational purposes one should, of course, attempt to find a gauge (potentials) that simplifies the computations as much as possible.

16.3 Motion in a Uniform Magnetic Field

The problem of the motion of an electron in a uniform magnetic field again became a "hot" topic in solid state physics in the 1980's due to the discovery by von Klitzing [16.2] of the *quantum Hall effect.* Much of the physics of this effect can be understood by solving the Schrödinger equation for an electron in a uniform magnetic field. The resultant discrete energy levels are called Landau levels [16.1] and are discussed in a later section. We begin by describing the integer quantum Hall effect. Then we solve the corresponding classical problem as well as the quantum problem. The integer quantum Hall effect manifests itself as a series of plateaus (fig. 16.1) in the Hall resistance R_H of materials in which the electrons are confined to motion in a plane. These are so-called two-dimensional electrons and occur at the interfacial oxide layer between a metal and a semiconductor in a device called a Metal-Oxide Semiconductor Field Effect Transistor or MOSFET. The actual layer is usually of the order of 100 Å so that the electrons are not truly two-dimensional but are confined to a thin layer.

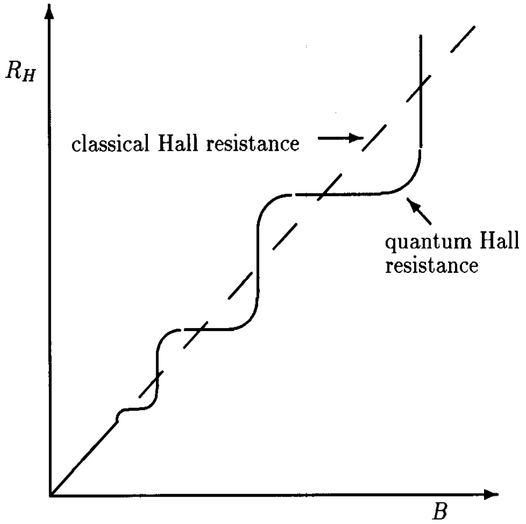

Figure 16.1: Schematic of plateaus in the quantum Hall effect.

16.3.1 Classical Hall Effect

To understand the term "Hall resistance" consider the schematic for an experiment shown in fig. 16.2. We assume that the current I flowing down the thin strip (thickness t and width w) is due to charge carriers of charge q and drift velocity v_d. In this case the current density $j = I/wt$ is given by

$$j = \frac{I}{wt} = Nqv_d \tag{16.3.30}$$

where N is the number of charge carriers per unit volume. Due to the magnetic field B the charge carriers experience a Lorentz force

$$F = \frac{1}{c} q v_d B \ . \tag{16.3.31}$$

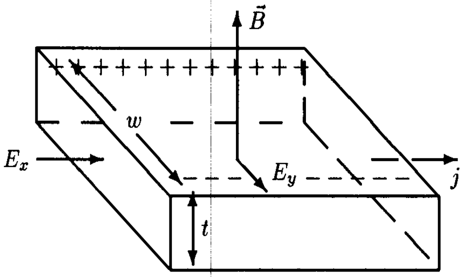

Figure 16.2: The Hall effect.

This force pushes the charges to one side of the strip, where they accumulate, and depletes the charges on the other side of the strip. This separation of charges produces an electric field

$$E = \frac{F}{q} = \frac{1}{c} v_d B \tag{16.3.32}$$

between the two sides of the strip and a force that just balances the Lorentz force as we have written down. This electric field leads to a potential difference, called the "Hall potential"

$$V_H = Ew \tag{16.3.33}$$

between the two sides of the strip.

The Hall resistance is defined by

$$R_H = \frac{V_H}{I} \ . \tag{16.3.34}$$

Using the results above, we find that

$$R_H = \frac{1}{c} \frac{B}{Ntq} \ . \tag{16.3.35}$$

It is convenient at this stage to introduce the two-dimensional or areal density of charge carriers

$$N_a = Nt \ . \tag{16.3.36}$$

Then we find that

$$R_H = \frac{1}{c}\frac{B}{N_a q} \quad . \tag{16.3.37}$$

It is worth noticing that the Hall resistance is positive if q is positive and negative if q is negative. Thus, an experiment in which one measure the Hall resistance, provides a means for measuring the sign of the charge on the charge carriers. Furthermore, the Hall resistance depends only on the areal density of charge carriers and no other properties of the underlying material.

16.3.2 Landau Levels

We now again solve the same problem, but this time we do it using quantum mechanics. We begin with the case of only a constant uniform magnetic field. The Hamiltonian is given by

$$H = \frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 \quad . \tag{16.3.38}$$

For convenience we choose the Landau gauge so that

$$\mathbf{A} = -B(y, 0, 0) \quad . \tag{16.3.39}$$

The Hamiltonian can now be written out explicitly and reads

$$H = \frac{1}{2m}\left(p_x - \frac{eB}{c}y\right)^2 + \frac{1}{2m}p_y^2 + \frac{1}{2m}p_z^2 \quad . \tag{16.3.40}$$

Since H does not contain any function of x or z we see that

$$[H, p_x] = [H, p_z] = 0 \quad . \tag{16.3.41}$$

Since p_x and p_z also commute we see that p_x, p_z and H form a complete set of observables and can be simultaneously diagonalized. To this end we look for a wavefunction of the form

$$\psi = \frac{1}{2\pi} e^{i(k_x x + k_z z)}\, \phi(y) \quad . \tag{16.3.42}$$

Then,

$$p_x \psi = \hbar k_x \psi \tag{16.3.43}$$

$$p_z \psi = \hbar k_z \psi \tag{16.3.44}$$

and the Schrödinger equation (16.3.34) reduces to

$$\frac{d^2\phi}{dy^2} + \left[\frac{2mE}{\hbar^2} - k_z^2 - \frac{m^2\omega^2}{\hbar^2}(y - y_0)^2\right]\phi = 0 \quad . \tag{16.3.45}$$

Here we have defined the Larmor frequency

$$\omega = \frac{eB}{mc} \tag{16.3.46}$$

and the parameter

$$y_0 = \hbar \frac{ck_x}{eB} \ . \tag{16.3.47}$$

Except for the constant k_z^2, equation (16.3.45) is just the equation for a simple harmonic oscillator centred at y_0. The energy is therefore given by

$$E_{n,k_z} = (n + 1/2)\hbar\omega + \frac{(\hbar k_z)^2}{2m} \ . \tag{16.3.48}$$

The corresponding wavefunction (up to a normalization constant A_n) is:

$$\psi_{n,k_x,k_z} = \frac{A_n}{2\pi} e^{i(k_x x + k_z z)} e^{-(y-y_0)^2/2\lambda^2} H_n((y - y_0)/\lambda) \ . \tag{16.3.49}$$

Here, as always, the H_n represent hermite polynomials and we have also introduced a new parameter

$$\lambda = \left(\frac{\hbar c}{eB}\right)^{1/2} \tag{16.3.50}$$

called the *magnetic length*. The physical significance of this parameter together with this problem are discussed in greater detail in the section after the next where we again solve this problem in a completely different manner.

It is worth noting that the energy does not depend on the wavenumber k_x. This means that each of the energy levels is infinitely degenerate. The physical reasons for this fact and some of its implications are also explored in the next sections. Furthermore, in the next section, we solve a mathematically very similar problem for crossed electric and magnetic fields. This, as we saw, is the situation of interest for the Hall effect.

16.4 Crossed Electric and Magnetic Fields

The Hall effect occurs, as we saw in the classical case, when uniform magnetic and electric fields perpendicular to each other are applied to a semiconducting medium. We model this effect by considering a single electron in empty space. The resultant Hall current is, as in the classical case, perpendicular to both the electric and magnetic fields.

Our model Hamiltonian is

$$H = \frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - e\phi \tag{16.4.51}$$

where we again choose the Landau gauge. Thus,

$$\mathbf{A} = -B(y, 0, 0) \quad , \quad \mathbf{B} = (0, 0, B) \tag{16.4.52}$$

and

$$\phi = -\mathcal{E}y \quad , \quad \mathbf{E} = (0, \mathcal{E}, 0) \ . \tag{16.4.53}$$

The Hamiltonian when written out reads

$$H = \frac{1}{2m}\left(p_x - \frac{eB}{c}y\right)^2 + \frac{1}{2m}p_y^2 + \frac{1}{2m}p_z^2 + e\mathcal{E}y \ . \tag{16.4.54}$$

Again we find that

$$[H, p_x] = [H, p_z] = 0 \tag{16.4.55}$$

and we can write, as before

$$\psi = \frac{1}{2\pi}\, e^{i(k_x x + k_z z)}\, \phi(y) \ . \tag{16.4.56}$$

The Schrödinger equation for $\phi(y)$ now reads

$$\frac{d^2\phi}{dy^2} - \frac{m^2\omega^2}{\hbar^2}\left[y - (y_0 - y_1)\right]^2 \phi + \left[\frac{2mE}{\hbar^2} - k_z^2 + \frac{m^2\omega^2}{\hbar^2} y_1 (2y_0 - y_1)\right]\phi = 0. \tag{16.4.57}$$

We have again introduced the parameters

$$y_0 = \hbar\frac{ck_x}{eB} \quad , \quad \omega = \frac{eB}{mc} \tag{16.4.58}$$

as well as the parameter

$$y_1 = \frac{e\mathcal{E}}{m\omega^2} \ . \tag{16.4.59}$$

Calling

$$a = y_0 - y_1 \tag{16.4.60}$$

we can now write the solution immediately as

$$\psi_{n,k_x,k_z} = \frac{A_n}{2\pi}\, e^{i(k_x x + k_z z)}\, e^{-(y-a)^2/2\lambda^2}\, H_n((y-a)/\lambda) \ . \tag{16.4.61}$$

The energy is given by

$$E_{n,k_x,k_z} = (n + 1/2)\hbar\omega + \frac{(\hbar k_z)^2}{2m} - \frac{m\omega^2}{2}\left(y_1^2 - 2y_1 y_0\right) \ . \tag{16.4.62}$$

The quantity of interest to us is the electric current density

$$\mathbf{J}_{n,k_x,k_z} = \frac{e\hbar}{2im}\left[\psi^* \left(\nabla + (ie/\hbar c)\mathbf{A}\right)\psi - \psi\left(\nabla - (ie/\hbar c)\mathbf{A}\right)\psi^*\right] \ . \tag{16.4.63}$$

The y-component of the current density clearly vanishes since the wavefunction describes a bound state in that direction. The other two components are given by

$$J_x = -\frac{e\hbar(k_x - y/\lambda^2)|\phi_n|^2}{(2\pi)^2 m} \tag{16.4.64}$$

$$J_z = -\frac{e\hbar k_z|\phi_n|^2}{(2\pi)^2 m} \ . \tag{16.4.65}$$

If we integrate over the y-coordinate to get the net current and use the fact that $|\phi_n|^2$ is an even function of y as well as that

$$\int_{-\infty}^{\infty} |\phi_n(y)|^2 \, dy = 1 \tag{16.4.66}$$

we find

$$\int_{-\infty}^{\infty} J_x(y) \, dy = -\frac{e\hbar k_x}{(2\pi)^2 m} \tag{16.4.67}$$

as well as

$$\int_{-\infty}^{\infty} J_z(y) \, dy = -\frac{e\hbar k_z}{(2\pi)^2 m} \quad . \tag{16.4.68}$$

If we now recall that $\hbar k_x$ and $\hbar k_z$ are just respectively the x and z-components of momentum then we see that these results coincide with the classical results.

16.4.1 The Quantum Hall Effect

We finally have all the ingredients for the quantum Hall effect. What we need is to compute the two-dimensional electron density N_a for the case of an electron in a magnetic field. This amounts to counting the degeneracy of the Landau levels for electrons confined to motion in a plane. Now the solution that we found, in section 16.3, for the wave function is

$$\psi_{n,k_x,k_z} = \frac{A_n}{2\pi} \, e^{i(k_x x + k_z z)} \, e^{-(y-a)^2/2\lambda^2} \, H_n((y-a)/\lambda) \, . \tag{16.4.69}$$

However, if the electrons are confined to the plane $z = 0$ then we see that the solution is of the form

$$\psi_{n,k_x,k_z} = \frac{A_n}{2\pi} \, e^{ik_x x} \, e^{-(y-a)^2/2\lambda^2} \, H_n((y-a)/\lambda) \, . \tag{16.4.70}$$

Here we have written a for y_0. Since the variable x is confined to $0 \le x \le L_x$ we need boundary conditions to make p_x self-adjoint. The simplest conditions are that the wavefunction be periodic. This means

$$\exp ik_x L_x = 1 \quad . \tag{16.4.71}$$

Therefore, we have for k_x the discrete values

$$k_r = \frac{2\pi r}{L_x} \qquad r = 1, 2, 3, \ldots \quad . \tag{16.4.72}$$

From equation (16.4.58) we recall that

$$a = y_0 = \hbar \frac{ck_x}{eB} \, . \tag{16.4.73}$$

This means that

$$k_x = \frac{eBy_0}{c\hbar} \le \frac{eBL_y}{c\hbar} = \frac{L_y}{\lambda^2} \quad . \tag{16.4.74}$$

Therefore, the index r has a maximum value

$$r_{\max} = \frac{k_{r\max} L_x}{2\pi} = \frac{L_x L_y}{2\pi\lambda^2} \; . \tag{16.4.75}$$

This number $r_{\max}$ is the number of states available in each Landau level (if we neglect spin) for a sample with area $L_x L_y$. Thus, $r_{\max}/(L_x L_y)$ determines the degeneracy per unit area of the sample in question. The resulting degeneracy is therefore

$$N_0 = \frac{r_{\max}}{L_x L_y} = \frac{eB}{hc} \; . \tag{16.4.76}$$

The considerations above suffice to obtain the Hall resistance. Each Landau level can hold the same number N_0 of electrons per unit area (neglecting spin). Therefore, if we have exactly j Landau levels filled then we must have

$$N_a = jN_0 \quad j = 1, 2, 3, \ldots \; . \tag{16.4.77}$$

Substituting this into the expression for the Hall resistance, namely

$$R_H = \frac{1}{c}\frac{B}{N_a e} \tag{16.4.78}$$

we get:

$$R_H = \frac{1}{c}\frac{B}{ejN_0} = \frac{h}{je^2} \quad j = 1, 2, 3, \ldots \; . \tag{16.4.79}$$

This simple equation illustrates why the quantum Hall effect is so important. The Hall resistance depends only the ratio of two fundamental constants, namely h and e^2. The Hall resistance provides the modern standard for resistance since

$$R_H = \frac{h}{je^2} = \frac{25\,812.81}{j} \; \Omega \; . \tag{16.4.80}$$

Furthermore, the quantum Hall effect also provides the most precise measurement of the value of Planck's constant h.

We have furthermore found that the discrete Landau levels explain the quantized steps in the Hall resistance. They do not, however, explain why the Hall resistance remains constant over a range of values of the magnetic field (the plateaus). This effect is due to defects in the material. These defects have the effect of spreading out the Landau levels. The discussion of how this happens is beyond the scope of this book, but is discussed in a qualitative fashion in the article by Halperin [16.2]. An argument based on gauge invariance is also to be found in the article by Laughlin [16.3].

16.5 Magnetic Field: Heisenberg Equations

We now return to the problem discussed in section 16.3 and solve it again. To illustrate how similar the classical equations of motion and the Heisenberg equations are, we solve the Heisenberg equations. To further illustrate the use of other gauges we this time employ the symmetric gauge so that with a uniform magnetic field in the z-direction we have

$$\mathbf{A} = \frac{B}{2}(-y, x, 0) \ . \tag{16.5.81}$$

Finally to obtain the wavefunctions we use an algebraic technique similar to the method used in section 9.2 to solve the harmonic oscillator. The Hamiltonian is again

$$H = \frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 \equiv \vec{\pi}^2 \tag{16.5.82}$$

where we have introduced the *mechanical momentum* $\vec{\pi}$ related to the *canonical momentum* $\mathbf{p}$ by

$$\vec{\pi} = \mathbf{p} + \frac{e}{c}\mathbf{A} \ . \tag{16.5.83}$$

If we were doing classical mechanics, the mechanical momentum $\vec{\pi}$ would just be equal to $m\mathbf{v}$. In quantum mechanics it is, however, the canonical momentum components p_j that satisfy the canonical commutation relations with the coordinate components x_k. In other words

$$[p_j, x_k] = -i\hbar\delta_{jk} \ . \tag{16.5.84}$$

Using this fact, and the explicit form for $\mathbf{A}$ given by (16.5.81) we can compute the various commutators among the mechanical momenta. Thus,

$$\begin{aligned} [\pi_x, \pi_y] &= \frac{e}{c}[p_x, A_y] + \frac{e}{c}[A_x, p_y] \\ &= \frac{eB}{2c}[p_x, x] + \frac{eB}{2c}[-y, p_y] \\ &= -\frac{ie\hbar}{c}B = -i\frac{\hbar^2}{\lambda^2} \ . \end{aligned} \tag{16.5.85}$$

Similarly, we find

$$[\pi_x, \pi_z] = [\pi_y, \pi_z] = 0 \ . \tag{16.5.86}$$

It is now straightforward to obtain the Heisenberg equations of motion

$$\dot{\pi}_x = \frac{1}{i\hbar}[\pi_x, H] = -\frac{eB}{mc}\pi_y = -\omega\pi_y \tag{16.5.87}$$

$$\dot{\pi}_y = \frac{1}{i\hbar}[\pi_y, H] = \frac{eB}{mc}\pi_x = \omega\pi_x \tag{16.5.88}$$

$$\dot{\pi}_z = \frac{1}{i\hbar}[\pi_z, H] = 0 \tag{16.5.89}$$

and

$$m\dot{\mathbf{r}} = \frac{m}{i\hbar}[\mathbf{r}, H] = \vec{\pi} \ . \tag{16.5.90}$$

Equation (16.5.89) shows that π_z is a constant of the motion. Furthermore writing out π_z explicitly we see that it coincides with the canonical momentum p_z. Thus,

$$\pi_z = p_z = \text{constant} . \tag{16.5.91}$$

This result is exactly the same as the result obtained earlier and also coincides with the classical result.

We next use equation (16.5.90) to replace π_y and π_x by $m\dot{y}$ and $m\dot{x}$ respectively on the right hand side of equations (16.5.87) and (16.5.88) to get

$$\dot{\pi}_x = -m\omega\dot{y} \tag{16.5.92}$$

$$\dot{\pi}_y = m\omega\dot{x} . \tag{16.5.93}$$

These equations can be integrated immediately to yield:

$$\pi_x + m\omega y = m\omega y_0 \tag{16.5.94}$$

$$\pi_y - m\omega x = -m\omega x_0 . \tag{16.5.95}$$

We have written the constants of integration as $m\omega y_0$ and $-m\omega x_0$. This means that x_0 and y_0 are also constants of the motion and commute with the Hamiltonian H. They also commute with both π_x and π_y. On the other hand they do not commute with each other. In fact

$$\begin{aligned} [x_0, y_0] &= \left[x - \frac{\pi_y}{m\omega}, y + \frac{\pi_x}{m\omega}\right] \\ &= -\frac{1}{m\omega}[\pi_y, y] + \frac{1}{m\omega}[x, \pi_x] - \frac{1}{m^2\omega^2}[\pi_y, \pi_x] . \end{aligned} \tag{16.5.96}$$

So, working this out, we find

$$[x_0, y_0] = \frac{i\hbar}{m\omega} = i\frac{\hbar c}{eB} = i\lambda^2 . \tag{16.5.97}$$

It is interesting to note that the equations of motion (16.5.87) to (16.5.90), as well as the first integrals (16.5.94) and (16.5.95) coincide with the classical results. For the classical motion we have free motion along the z-direction and circular motion in the $x - y$ plane with the centres of the circles located at (x_0, y_0). In fact, for constant energy we have

$$H = \frac{1}{2m}\left(\pi_x^2 + \pi_y^2\right) + \frac{p_z^2}{2m} = E_t + \frac{p_z^2}{2m} \tag{16.5.98}$$

where we have introduced the conserved "transverse energy" E_t. If we now consider $(2E_t)/(m\omega^2)$ and write this out we get *classically*

$$\frac{2E_t}{m\omega^2} = \frac{1}{m^2\omega^2}\left(\pi_x^2 + \pi_y^2\right) = (x - x_0)^2 + (y - y_0)^2 . \tag{16.5.99}$$

This is the equation for a circular orbit with radius

$$R^2 = \frac{2E_t}{m\omega^2} \tag{16.5.100}$$

and centre at (x_0, y_0). The same result holds quantum mechanically. This does not mean, however, that the trajectory of such an electron in a fixed energy eigenstate of the Hamiltonian

$$H_t = \frac{1}{2m}\left(\pi_x^2 + \pi_y^2\right) \tag{16.5.101}$$

is exactly measurable. Only the radius R is measurable. The centre of the circle is uncertain because x_0 and y_0 do not commute but satisfy instead the commutation relation given by equation (16.5.97). Thus, they also satisfy the uncertainty relation

$$\Delta x_0 \Delta y_0 \geq \frac{\lambda^2}{2} \,. \tag{16.5.102}$$

So if, in conformity with the classical results, we interpret (x_0, y_0) as the operators whose eigenvalues yield the coordinates of the centre of the circular motion, then the centre of the circle is not exactly measurable. This is why a trajectory is not observable. Writing out the Hamiltonian (16.5.101), we find

$$\begin{aligned} H_t &= \frac{1}{2m}\left(p_x^2 + p_y^2\right) + \frac{e^2B^2}{8mc^2}(x^2 + y^2) + \frac{eB}{2mc}(xp_y - yp_x) \\ &= \frac{1}{2m}\left(p_x^2 + p_y^2\right) + \frac{1}{8}m\omega^2(x^2 + y^2) + \frac{1}{2}\omega L_z \,. \end{aligned} \tag{16.5.103}$$

We can easily check that L_z is also a constant of the motion since it commutes with both H_t and $p_z^2/2m$. It is, however, not an independent constant of the motion since we can write

$$L_z = xp_y - yp_x = x\pi_y - y\pi_x - \frac{1}{2}m\omega(x^2 + y^2) \tag{16.5.104}$$

and then use equations (16.5.94), (16.5.95) and (16.5.99) to rewrite this as

$$\begin{aligned} L_z &= m\omega[x(x - x_0) + y(y - y_0) - \frac{1}{2}(x^2 + y^2)] \\ &= \frac{1}{2}m\omega(R^2 - x_0^2 - y_0^2) \,. \end{aligned} \tag{16.5.105}$$

This concludes our discussion of constants of the motion and solutions of the Heisenberg equations. The physical interpretation is now clear. The motion of the electron is as in the classical case, there is linear motion parallel to the magnetic field and circular motion about the magnetic field. Only the location of the centre of the orbit is indeterminate to roughly within a circle of radius λ.

16.6 Energy Eigenfunctions

We can also obtain the energy eigenfunctions using the Heisenberg operators. In fact the procedure is similar to what we did for the simple harmonic oscillator when we solved that problem algebraically. However, because there are additional conserved quantum numbers the problem is somewhat more complicated. To start we introduce, in the *symmetric gauge*, the operators

$$a = \frac{\lambda}{\hbar\sqrt{2}}(\pi_y + i\pi_x) = \frac{1}{\lambda\sqrt{2}}\left[\frac{x - iy}{2} + \lambda^2\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\right] \tag{16.6.106}$$

$$a^\dagger = \frac{\lambda}{\hbar\sqrt{2}}(\pi_y - i\pi_x) = \frac{1}{\lambda\sqrt{2}}\left[\frac{x + iy}{2} - \lambda^2\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\right] . \tag{16.6.107}$$

Then,

$$[a, a^\dagger] = 1 \tag{16.6.108}$$

and the Hamiltonian H_t can be written as

$$H_t = \hbar\omega(a^\dagger a + 1/2) \ . \tag{16.6.109}$$

So we have exactly recovered the Hamiltonian for a simple harmonic oscillator. This means we can define the ground state by

$$a\,\phi_0 = 0 \tag{16.6.110}$$

and

$$\pi_z\,\phi_0 = \hbar k\,\phi_0 \ . \tag{16.6.111}$$

The normalized state ϕ_n is then given by

$$\phi_n = \frac{1}{\sqrt{n!}}\left(a^\dagger\right)^n \phi_0 \ . \tag{16.6.112}$$

To find ϕ_0 we use the explicit form of a as given by eqn (16.6.106) as well as the explicit form of $\pi_z = -i\hbar\partial/\partial z$ to solve equations (16.6.110) and (16.6.111). We then obtain

$$\phi_{0,f} = f(x - iy)\exp\left\{ikz - \frac{x^2 + y^2}{4\lambda^2}\right\} \tag{16.6.113}$$

where f is an arbitrary function. This fact reflects the infinite degeneracy of these states. It also means that we can impose another condition on ϕ_0. A possible condition is that ϕ_0 should also be an eigenfunction of either x_0 or y_0. Either of these conditions is motivated by the fact that

$$[x_0, a] = [x_0, a^\dagger] = [y_0, a] = [y_0, a^\dagger] = 0 \tag{16.6.114}$$

and thus if

$$x_0\phi_0 = \xi\,\phi_0 \tag{16.6.115}$$

or

$$y_0\phi_0 = \eta\,\phi_0 \ , \tag{16.6.116}$$

we see, if we use (16.6.112), that we also have

$$x_0\phi_{n,\xi} = \xi\,\phi_{n,\xi} \tag{16.6.117}$$

or

$$y_0\phi_{n,\eta} = \eta\,\phi_{n,\eta} \,. \tag{16.6.118}$$

Writing out x_0 and y_0 we find

$$x_0 = \frac{x}{2} + i\lambda^2\frac{\partial}{\partial y} \tag{16.6.119}$$

$$y_0 = \frac{y}{2} - i\lambda^2\frac{\partial}{\partial x} \,. \tag{16.6.120}$$

The use of either (16.6.117) or (16.6.118) destroys the obvious symmetry between x and y. This suggests that we introduce the more symmetric operators

$$b = \frac{x_0 + iy_0}{\lambda\sqrt{2}} = \frac{1}{\lambda\sqrt{2}}\left[\frac{x+iy}{2} + \lambda^2\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\right] \tag{16.6.121}$$

$$b^\dagger = \frac{x_0 - iy_0}{\lambda\sqrt{2}} = \frac{1}{\lambda\sqrt{2}}\left[\frac{x-iy}{2} - \lambda^2\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\right] \,. \tag{16.6.122}$$

Then we find that

$$[b, b^\dagger] = 1 \quad \text{and} \quad [a, b] = [a, b^\dagger] = 0 \,. \tag{16.6.123}$$

So both b and $b^\dagger$ commute with the Hamiltonian H_t. This means that we can label the eigenstates not only by the Landau level quantum number n but also by the eigenvalue m of $b^\dagger b$. That is, we solve instead of either equation (16.6.117) or (16.6.118) the equation

$$b^\dagger b\,\phi_{n,m} = m\,\phi_{n,m} \,. \tag{16.6.124}$$

This is easily achieved by setting

$$b\,\phi_{n,0} = 0 \tag{16.6.125}$$

and

$$\phi_{n,m} = \frac{1}{\sqrt{m!}}\left(b^\dagger\right)^m \phi_{n,0} \,. \tag{16.6.126}$$

The explicit solution for $\phi_{0,0}$ is given by:

$$\phi_{0,0} = \frac{1}{\sqrt{2\pi\lambda^2}}\exp\left\{ikz - \frac{x^2+y^2}{4\lambda^2}\right\} \,. \tag{16.6.127}$$

16.7 Translation Invariant States

We now use the results just obtained to construct translation invariant states. In doing so we follow closely the paper by Ferrari [16.4]. Such states are again solutions of the Schrödinger equation with the Hamiltonian H_t but, in addition, they also satisfy boundary conditions on the boundary of the material. We specify the precise form of these boundary conditions after we first develop some machinery. To this end we consider the operator

$$S(c) = \exp[\frac{c^*b - cb^\dagger}{\sqrt{2}}] \tag{16.7.128}$$

where c is a complex number related to a two-dimensional vector $\mathbf{c}$ with components

$$\mathbf{c} = (c_x, c_y, 0) \tag{16.7.129}$$

by

$$c = c_x + ic_y \ . \tag{16.7.130}$$

The operator $S(c)$ may also be written

$$S(c) = \exp[\frac{i}{\hbar}\mathbf{c} \cdot (\mathbf{p} - e/c\mathbf{A})] \ . \tag{16.7.131}$$

It is important to notice the sign in front of the vector potential; it is the *opposite* of the sign appearing in the term defining the mechanical momentum operator $\vec{\pi}$. The operator $S(c)$ is known as a *coherent-state* operator and as such has been extensively studied. For further references see the paper by Ferrari [16.4]. Now suppose that d is another complex number corresponding to a vector $\mathbf{d}$. Then,

$$S(c)S(d) = S(c+d)\exp[-\frac{c^*d - cd^*}{4\lambda^2}] \ . \tag{16.7.132}$$

The argument of the exponential when written out reads

$$-\frac{c^*d - cd^*}{4\lambda^2} = -\frac{i}{2\lambda^2}(c_y d_x - c_x d_y) = -\frac{i}{2\lambda^2}(\mathbf{c} \times \mathbf{d}) \cdot \mathbf{e}_z \ . \tag{16.7.133}$$

Hence, we find that

$$S(c)S(d) = S(c+d)\exp[-\frac{i}{2\lambda^2}(\mathbf{c} \times \mathbf{d}) \cdot \mathbf{e}_z] \tag{16.7.134}$$

and, repeating the process

$$S(c)S(d) = S(d)S(c)\exp[-\frac{i}{\lambda^2}(\mathbf{c} \times \mathbf{d}) \cdot \mathbf{e}_z] \ . \tag{16.7.135}$$

Therefore, if the parallelogram with sides given by the vectors $\mathbf{c}$ and $\mathbf{d}$ has an area which is an integral multiple of $2\pi\lambda^2$, that is,

$$(\mathbf{c} \times \mathbf{d}) \cdot \mathbf{e}_z = 2\pi\lambda^2 u \quad u = 1, 2, 3, \ldots \tag{16.7.136}$$

then $S(c)$ and $S(d)$ commute

$$[S(c), S(d)] = 0 \ . \tag{16.7.137}$$

So, remembering that $\lambda^2 = (\hbar c)/(eB)$, we see that u is the number of units of flux, of size hc/e, passing through the parallelogram with sides $\mathbf{c}$, $\mathbf{d}$. From now on we shall only consider vectors $\mathbf{c}$, $\mathbf{d}$ such that (16.7.136) is satisfied. Clearly the operators $S(c)$ commute with both a and $a^\dagger$ and hence with the Hamiltonian H_t. The action of $S(c)$ on coordinates is obtained just as for the case of the squeezing operator in section 11.5. One begins by seeing that if we write $S(c)$ as

$$S(c) = e^F \tag{16.7.138}$$

then

$$F = \frac{i}{\hbar}\mathbf{c}\cdot(\mathbf{p} - e/c\,\mathbf{A}) \tag{16.7.139}$$

and

$$[\mathbf{r}, F] = -\mathbf{c} \ . \tag{16.7.140}$$

Therefore, it follows that

$$\mathbf{r}\, e^{-F} = e^{-F}(\mathbf{x} + \mathbf{c}) \ , \tag{16.7.141}$$

so that

$$S(c)\mathbf{r}S(c)^{-1} = S(c)\mathbf{r}S(-c) = \mathbf{r} + \mathbf{c} \ . \tag{16.7.142}$$

This does not mean that $S(c)$ is simply a translation operator. It is a translation operator followed by a coordinate-dependent phase as is clear when it is rewritten as

$$S(c) = \exp\left[\frac{i}{2\lambda^2}(\mathbf{c}\times\mathbf{r})\cdot\mathbf{e}_z\right]\,\exp\left[c_x\frac{\partial}{\partial x} + c_y\frac{\partial}{\partial y}\right] \ . \tag{16.7.143}$$

Since the range of x and y is finite, π_x and π_y have each a one-parameter family of self-adjoint extensions. These extensions may be specified by the boundary conditions

$$\begin{aligned} S(L_1)\psi &= e^{i\theta_1}\psi \\ S(L_2)\psi &= e^{i\theta_2}\psi \end{aligned} \tag{16.7.144}$$

where θ_1, θ_2 label these unitarily inequivalent self-adjoint extensions of the coherent state operator $\mathbf{p} - e/c\mathbf{A}$. In our previous discussion of the quantum Hall effect we had chosen the parameters $\theta_1 = \theta_2 = 0$. We now look at what parameters are possible. The conditions (16.7.144) are clearly compatible only if

$$S(L_1)S(L_2) = S(L_2)S(L_1) \tag{16.7.145}$$

so that we require

$$|(\mathbf{L}_1 \times \mathbf{L}_2) \cdot \mathbf{e}_z| = 2\pi\lambda^2 g_L \quad g_L = 1,2,3,\ldots \ . \tag{16.7.146}$$

It will turn out that g_L is the degeneracy of the Landau level, just as we saw before. Hence the subscript L. Now if

$$\mathbf{L}_1 = p\,\mathbf{c} \quad p = 1,2,3,\ldots \tag{16.7.147}$$

and

$$\mathbf{L}_2 = p'\,\mathbf{c} + q\mathbf{d} \quad p'\,,\, q = 1,2,3,\ldots \tag{16.7.148}$$

with as above

$$(\mathbf{c} \times \mathbf{d}) \cdot \mathbf{e}_z = 2\pi\lambda^2 u \quad u = 1,2,3,\ldots \tag{16.7.149}$$

we find that

$$g_L = pqu \tag{16.7.150}$$

and the vectors **c**, **d** form a lattice which provides a tiling of the area A bounded by $\mathbf{L}_1$ and $\mathbf{L}_2$.

We now use these operators $S(c)$ and $S(d)$ to construct a translation covariant basis. To do this we begin with the wavefunction $\phi_{n,0}$ in equation (16.6.125) and form the double series

$$\psi_n^{\mu,\nu}(\mathbf{r}) = \frac{1}{\sqrt{pq}} \sum_{r,s=-\infty}^{\infty} \left[S(c)\, e^{-i\mu}\right]^r \left[S(d)\, e^{-i\nu}\right]^s \phi_{n,0}(\mathbf{r}) \ . \tag{16.7.151}$$

Using the explicit form of the operator S we can rewrite this expression as

$$\begin{aligned}&\psi_n^{\mu,\nu}(\mathbf{r})\\ &= \frac{1}{\sqrt{pq}} \sum_{r,s=-\infty}^{\infty} (-1)^{rsu}\, e^{-i(\mu r+\nu s)}\, e^{i/2\mathbf{e}_z\cdot(r\mathbf{c}+s\mathbf{d})\times\mathbf{r}}\, \phi_{n,0}(\mathbf{r} + r\mathbf{c} + s\mathbf{d}) \ .\end{aligned} \tag{16.7.152}$$

This series converges pointwise since $\phi_{n,0}$ is Gaussian for large values of the argument. Furthermore, since the operators $S(c)$ and $S(d)$ both commute with H_t this is again an eigenstate of H_t with eigenvalue $(n+1/2)\hbar\omega$. This function is a Bloch function constructed from the coherent states associated with the lattice formed by the vectors **c** and **d**. The phases μ and ν are fixed by the boundary conditions (16.7.144). In fact we find

$$\begin{aligned} S(c)\psi_n^{\mu,\nu} &= e^{i\mu}\, \psi_n^{\mu,\nu} \\ S(d)\psi_n^{\mu,\nu} &= e^{i\nu}\, \psi_n^{\mu,\nu} \ . \end{aligned} \tag{16.7.153}$$

Now, using the fact that

$$S(L_1) = S(pc) = (S(c))^p \tag{16.7.154}$$

and

$$S(L_2) = S(p'c + qd) = S(p'c)\,S(qd)\exp\left\{\frac{ip'q}{2\lambda^2}(\mathbf{c}\times\mathbf{d})\cdot\mathbf{e}_z\right\} \tag{16.7.155}$$

we find that the boundary conditions are satisfied if

$$\begin{aligned} p\mu &= \theta_1 + 2\pi n_1 \\ \pi u p' q + p'\mu + q\nu &= \theta_2 + 2\pi n_2 \end{aligned} \tag{16.7.156}$$

or

$$\begin{aligned} \mu &= = \frac{1}{p}[\theta_1 + 2\pi n_1] \\ \nu &= \frac{2\pi}{pq}[-p'n_1 + pn_2] + \frac{1}{pq}[-p'\theta_1 + p\theta_2] \end{aligned} \tag{16.7.157}$$

where

$$\begin{aligned} &n_1 = 0, 1, \ldots, p-1 \\ &0 \le ([-p'n_1 + pn_2) \le pq \ . \end{aligned} \tag{16.7.158}$$

Thus, we have a set of eigenfunctions of H_t that also satisfy the boundary conditions (16.7.144). Clearly, different choices of θ_1 and θ_2 correspond to different self-adjoint extensions of the coherent state operator $\mathbf{p} - e/c\mathbf{A}$ and hence to different dynamics. The difficult questions of normalization and other properties of these solutions as well as further references to the literature are to be found in the paper by Ferrari [16.4].

16.8 Gauge Transformations

The calculations so far were performed in the symmetric gauge. This was done only for convenience. We now show that the whole procedure goes through for any gauge. To do this we again write equations (16.6.106) and (16.6.107) as

$$a = \frac{\lambda}{\hbar\sqrt{2}}(\pi_y + i\pi_x) \tag{16.8.159}$$

$$a^\dagger = \frac{\lambda}{\hbar\sqrt{2}}(\pi_y - i\pi_x) \tag{16.8.160}$$

where, as before the mechanical momentum is

$$\vec{\pi} = \mathbf{p} + \frac{e}{c}\mathbf{A} \ . \tag{16.8.161}$$

However, this time the vector potential $\mathbf{A}$ is in an arbitrary gauge. Furthermore, we also define

$$\begin{aligned} b &= -a^\dagger + \frac{x + iy}{\lambda\sqrt{2}} \\ b^\dagger &= -a + \frac{x - iy}{\lambda\sqrt{2}} \ . \end{aligned} \tag{16.8.162}$$

This reduces to our previous definition (16.6.121) and (16.6.122) for the symmetric gauge. It now again follows that

$$[a,a^\dagger]=1 \quad , \quad [b,b^\dagger]=1 \quad , \quad [a,b]=[a,b^\dagger]=0 \, . \tag{16.8.163}$$

So these operators satisfy the same algebra as those previously defined and hence are related to the old operators by a unitary transformation. This transformation amounts to multiplying each state by a phase factor. Thus, if we let

$$\mathbf{A}=\frac{B}{2}(-y,x,0)+\nabla\Lambda(x,y) \tag{16.8.164}$$

then the unitary operator is given by

$$U=\exp\left(-\frac{ie}{\hbar c}\Lambda\right) \, . \tag{16.8.165}$$

Furthermore, the boundary conditions do not change since the transformation given is unitary.

For example, if we choose

$$\Lambda=\frac{B}{2}xy \tag{16.8.166}$$

we recover for the basis functions, the solutions in the Landau gauge (equation (16.3.49)). Also we find that the factor $\exp(ik_x x)$ is replaced by a factor $\exp[(iy(x-2a))/(2\lambda^2)]$. Now, using the fact that $a=x_0=-(\hbar c)/(eB)k_y$ and $\lambda^2=(\hbar c)/(eB)$ we find that this factor is just $\exp[i(xy)/(2\lambda^2)+ik_y y]$. The first factor is simply due to the gauge transformation from the Landau to the symmetric gauge with the gauge function

$$\Lambda=\frac{B}{2}xy \, . \tag{16.8.167}$$

The factor

$$\exp[-ik_x x+ik_y y]=\exp[i(xy_0-yx_0)/(2\lambda^2)] \tag{16.8.168}$$

reflects the arbitrariness in the choice of the centre $(x_0,\, y_0)$ for the orbit.

These eigenfunctions not only play an important role in understanding the quantum Hall effect but also the energy levels of atoms in superstrong magnetic fields. By superstrong magnetic fields we mean fields so strong that the "orbits" due to the magnetic field have shrunk to the order of magnitude of Bohr orbits. Thus, these magnetic fields must be so strong that their Larmor orbit

$$R_0=\sqrt{\frac{\hbar\omega}{m\omega^2}}\le a_0=\frac{\hbar^2}{me^2} \, . \tag{16.8.169}$$

Substituting for the Larmor frequency ω and solving for B we find

$$B\ge\frac{m^2e^3}{\hbar^3}c \, . \tag{16.8.170}$$

For such strong fields it is not the magnetic interaction that is the perturbation but rather the Coulomb term [16.2] and the atoms become more "cigar" than "spherical" in shape.

16.9 Problems

16.1 The strongest static magnetic fields currently achieved in laboratories are of the order of 3×10^5 gauss. For fields of this strength estimate the magnitude of the term

$$\frac{e^2}{2mc^2}\mathbf{A}\cdot\mathbf{A}\ .$$

16.2 Solve the eigenvalue problem for

$$r_0^2 = x_0^2 + y_0^2$$

where x_0 and y_0 are given by equations (16.4.70) and (16.4.71). It may be useful to write the eigenvalue problem in the form

$$r_0^2 f_l = (2l+1)\lambda^2 f_l\ .$$

Interpret the meaning of this result.

16.3 Show that in the presence of a time-dependent electromagnetic field $(\phi, \mathbf{A})$ the equation of continuity holds for a particle of charge $-e$ if the charge density is given by

$$\rho = -e\,\Psi^*\Psi$$

and the current density is given by

$$\mathbf{j} = -e\frac{\hbar}{2im}\left[\Psi^*(\nabla + ie/(\hbar c)\mathbf{A})\Psi - \Psi(\nabla - ie/(\hbar c)\mathbf{A})\Psi^*\right]\ .$$

Hint: Start with the time-dependent Schrödinger equation.

16.4 Show that the Schrödinger equation is form invariant under the gauge transformations given by equations (16.2.21).

16.5 Under a gauge transformation we have that for static electromagnetic fields

$$\mathbf{A} \to \mathbf{A}^\Lambda = \mathbf{A} + \nabla\Lambda$$

$$\phi \to \phi^\Lambda = \phi$$

$$\Psi \to \Psi^\Lambda = e^{i(q/\hbar c)\Lambda}\,\Lambda$$

so that the Schrödinger equation remains form invariant. We also require, however, that observables be gauge invariant in the sense that their matrix elements remain invariant under gauge transformations. This means that if under a gauge transformation an observable

$$O(\mathbf{p},\mathbf{A},\phi) \to O(\mathbf{p},\mathbf{A}^{\Lambda},\phi^{\Lambda})$$

we require that

$$(\Psi, O(\mathbf{p},\mathbf{A},\phi)\Phi) = (\Psi^{\Lambda}, O(\mathbf{p},\mathbf{A}^{\Lambda},\phi^{\Lambda})\Phi^{\Lambda}) \;.$$

The left side may be rewritten as

$$(\Psi, O(\mathbf{p},\mathbf{A},\phi)\Phi) = (\Psi^{\Lambda}, O^{\lambda}(\mathbf{p},\mathbf{A},\phi)\Phi^{\Lambda}) \;,$$

where

$$O^{\lambda}(\mathbf{p},\mathbf{A},\phi) = e^{i(q/\hbar c)\Lambda}\, O(\mathbf{p},\mathbf{A},\phi)\, e^{-i(q/\hbar c)\Lambda}$$

is the unitary transform of the operator $O(\mathbf{p},\mathbf{A},\phi)$. Show that the resulting necessary and sufficient condition, namely

$$O^{\lambda}(\mathbf{p},\mathbf{A},\phi) = O(\mathbf{p},\mathbf{A}^{\Lambda},\phi^{\Lambda})$$

is satisfied if and only if

$$O^{\lambda}(\mathbf{p},\mathbf{A},\phi) = O(\mathbf{p}+q/c\mathbf{A},\phi) \;\;.$$

See reference [16.5] for further discussion.

Bibliography

[16.1] L. Landau, Zeitschrift fr Physik, **64**, 629, (1930).

[16.2] The original paper on the quantum Hall effect is:
K. von Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).
von Klitzing's Nobel lecture is published in:
K. von Klitzing, Rev. Mod. Phys. **58**, 519 (1986).
See also the Scientific American article: B. I. Halperin, Sci. Am. **254**, 52 (1986).

[16.3] R.B. Laughlin, Phys. Rev. **B23**, 5632 (1981).

[16.4] R. Ferrari, Phys. Rev. **B42**, 4598, (1990).

[16.5] D.H. Kobe, Am. J. Phys. **54**, 77, (1986).

Chapter 17

Applications

17.1 Introduction

In this chapter we continue the study of the interaction of electrons with the electromagnetic field. In particular, we study the splitting of spectral lines due to a magnetic field - the Zeeman effect. To carry out computations for this effect requires the study of the addition of angular momenta. This problem constitutes a major portion of this chapter.

17.2 Spin and Spin-Orbit Coupling

We now consider an electron interacting with both an electric as well as a constant magnetic field. The "orbital" part of the Hamiltonian is then given by (16.2.3) with $q = -e$,

$$H_o = \frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - e\phi \tag{17.2.1}$$

where the subscript o stands for "orbital". Choosing the symmetric gauge such that

$$\mathbf{A} = -\frac{1}{2}\mathbf{r} \times \mathbf{B} \tag{17.2.2}$$

we can rewrite this Hamiltonian (see problem 17.8) as

$$H_o = \frac{\mathbf{p}^2}{2m} + V(r) + \frac{e}{2mc}\mathbf{B} \cdot \mathbf{L} + \frac{e^2}{2mc^2}\mathbf{A}^2 - e\phi \tag{17.2.3}$$

where $V = -e\phi$. For most cases of weak magnetic fields we drop the second last term since it is much smaller than the other terms.

This is not yet the complete Hamiltonian for an electron in an electromagnetic field. The fact that the electron has spin adds two more terms to this

Hamiltonian. The first of these arises from the fact that due to its spin the electron also has a magnetic moment (section 9.7)

$$\vec{\mu} = \frac{ge}{2mc}\mathbf{S} \, . \tag{17.2.4}$$

The factor g is known as the *gyromagnetic ratio* and was first measured spectroscopically to have the value $g = -2$. This factor arises automatically in Dirac's relativistic equation for an electron and is exactly -2 except for very small corrections due to quantum electrodynamic effects. The measured value of g agrees with the value calculated from quantum electrodynamics and is

$$g = -2 \times (1 + 1.159\,638\,9 \times 10^{-3}) \, . \tag{17.2.5}$$

We shall use the value $g = -2$. Thus, the energy due to this magnetic moment is given by

$$H_m = -\vec{\mu} \cdot \mathbf{B} = -\frac{ge}{2mc}\mathbf{B} \cdot \mathbf{S} \, . \tag{17.2.6}$$

There is a second energy term which arises due to the interaction of the spin magnetic moment of the electron with the magnetic field created by its orbital motion. Crudely speaking, the electromagnetic field due to the nucleus as described in (17.2.6) above is viewed in the rest frame of the nucleus. However, to get the same field in the rest frame of the electron we must perform a Lorentz transformation. This causes the originally purely electric field

$$\mathbf{E} = -\frac{\partial \phi}{\partial r}\frac{\mathbf{r}}{r} \tag{17.2.7}$$

to acquire a magnetic component

$$\mathbf{B}_{\text{Lorentz}} = -\frac{\mathbf{v}}{c} \times \mathbf{E} \, . \tag{17.2.8}$$

The interaction of this magnetic field gives rise to the so-called spin-orbit interaction

$$H'_{s.o.} = -\vec{\mu} \cdot \mathbf{B}_{\text{Lorentz}} = -\frac{ge}{2mc}\left(\frac{\mathbf{v}}{c} \times \mathbf{E}\right) \cdot \mathbf{S} \tag{17.2.9}$$

or

$$H'_{s.o.} = \frac{e}{m^2c^2}\mathbf{S} \cdot (\mathbf{E} \times \mathbf{p}) \, . \tag{17.2.10}$$

This term was first proposed by Goudsmit and Uehlenbeck [17.1]. Unfortunately it was too large by a factor of 2. Thomas, however, showed that relativistic effects cause a further precession of the spin vector to effectively reduce $H'_{s.o.}$ by a factor of 2. Thus, the spin-orbit interaction is given by

$$H_{s.o.} = \frac{e}{2m^2c^2}\mathbf{S} \cdot (\mathbf{E} \times \mathbf{p}) \, . \tag{17.2.11}$$

If the central electrostatic field is described by a potential $\phi(r)$ then

$$\mathbf{E} = -|\nabla\phi|\frac{\mathbf{r}}{r} = -\frac{1}{r}\frac{d\phi}{dr}\mathbf{r} = \frac{1}{er}\frac{dV}{dr}\mathbf{r} \tag{17.2.12}$$

where we have introduced the potential energy $V = -e\phi$. So finally

$$H_{s.o.} = \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\mathbf{S}\cdot\mathbf{L}\,. \tag{17.2.13}$$

Thus, the resultant Hamiltonian is given by

$$H = \frac{\mathbf{p}^2}{2m} + V(r) + \frac{e}{2mc}\mathbf{B}\cdot(\mathbf{L}+2\mathbf{S}) + \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\mathbf{S}\cdot\mathbf{L}\,. \tag{17.2.14}$$

17.3 Alkali Spectra

Consider any alkali atom in the absence of a magnetic field. Furthermore restrict your attention to the outer or valence electrons and treat the inner electrons as inert closed shells. Then this is effectively a one electron problem in some central electrostatic potential $V(r)$. Then, the appropriate corresponding Hamiltonian is given by (17.2.14) with $\mathbf{B} = 0$

$$H = \frac{\mathbf{p}^2}{2m} + V(r) + \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\mathbf{S}\cdot\mathbf{L}\ . \tag{17.3.15}$$

Now, the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ satisfies

$$\mathbf{J}^2 = \mathbf{L}^2 + \mathbf{S}^2 + 2\mathbf{L}\cdot\mathbf{S} \tag{17.3.16}$$

so that

$$\mathbf{L}\cdot\mathbf{S} = \frac{1}{2}\left(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2\right)\ . \tag{17.3.17}$$

Also since H, $\mathbf{J}^2$, $\mathbf{L}^2$ and $\mathbf{S}^2$ are mutually commuting we can label the eigenkets of H by $|n,j,l,s\rangle$. Thus, fixing j and l the effective potential becomes

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2}{4m^2c^2}\frac{1}{r}\frac{dV}{dr}[j(j+1) - l(l+1) - 3/4]\,. \tag{17.3.18}$$

Here we have used the fact that $S = \hbar/2$. Thus, in fact, $j = l \pm 1/2$ except for S-waves ($l = 0$) in which case $j = 1/2$. The effect of this is to give a different effective potential for levels with the same n and l but different orientations of $\mathbf{L}$ and $\mathbf{S}$ or in fact different $\mathbf{J}$, except for S-waves. This means that all degenerate levels with $l \neq 0$ have in fact two different effective potentials depending on the two values of j. So, all energy levels with $l \neq 0$ are doublets. For example, all $l = 1$ or P levels split into the doublets $P_{3/2}$ and $P_{1/2}$ where the subscript ers to the value of j.

If we now treat the spin-orbit term in (17.3.15) as a perturbation then to first order in perturbation theory we can calculate the splitting resulting from the spin-orbit interaction. Thus,

$$\begin{aligned}\Delta E_{\text{doublet}} &= E_{n,l,l+1/2} - E_{n,l,l-1/2}\\ &= \frac{\hbar^2}{4m^2c^2}\langle n,l|\frac{1}{r}\frac{dV}{dr}|n,l\rangle\\ &\quad [(l+1/2)(l+3/2) - l(l+1) - 3/4\\ &\quad -(l-1/2)(l+1/2) + l(l+1) + 3/4]\end{aligned} \tag{17.3.19}$$

or

$$\Delta E_{\text{doublet}} = \frac{\hbar^2}{4m^2c^2}\langle n,l|\frac{1}{r}\frac{dV}{dr}|n,l\rangle(2l+1) \tag{17.3.20}$$

where we have written $|n,l\rangle$ for both $|n,l+1/2,l,1/2\rangle$, and $|n,l-1/2,l,1/2\rangle$ since the matrix elements for both these states are identical.

It is this interaction that is responsible for the well known spectroscopic doublet in sodium called the sodium D lines. This doublet occurs in the transition from the $P_{3/2}$ and $P_{1/2}$ states to S states. Here we are using the spectroscopic notation S, P, D, F, G, H, etc. for $l = 0,1,2,3,4,5$ etc. and the subscript denotes the j value of the level.

The reason that the sodium doublet is more widely split than the hydrogen doublet is due to the more rapid variation of $V(r)$. Thus, dV/dr is much larger for sodium than hydrogen since the inner closed shells in the case of sodium screen the valence electron and more of the nuclear charge becomes effective on the valence electron as it penetrates the outer shell.

After considering spin-orbit coupling the obvious next step is to consider an atom in a weak magnetic field as well and to study the Hamiltonian (17.2.14). Unfortunately we must first study another aspect of angular momentum, namely how to add two angular momenta since (17.2.14) involves the term $\mathbf{L}+\mathbf{S}$. We shall therefore do this first and then return to the Hamiltonian given by (17.2.14).

17.4 Addition of Angular Momenta

Consider a pair of particles with angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$. Then,

$$[\mathbf{J}_1,\mathbf{J}_2] = 0\ . \tag{17.4.21}$$

This will also be the case if $\mathbf{J}_1$ refers to the orbital angular momentum $\mathbf{L}$ and $\mathbf{J}_2$ refers to the spin angular momentum $\mathbf{S}$ of just one particle. In either case the total angular momentum $\mathbf{J}$ is given by

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2\ . \tag{17.4.22}$$

If we now consider commutators, again picking z as the preferred direction, we find two sets of mutually commuting operators, namely

1) J^2, J_z, J_1^2, J_2^2

and

2) J_1^2, J_{1z}, J_2^2, J_{2z} .

Thus, it is possible to simultaneously diagonalize all the operators in either the first or the second set. We label the corresponding eigenkets as

1) $|j,m,j_1,j_2\rangle$

and

2) $|j_1, m_1, j_2, m_2\rangle$.

Both sets of eigenkets then form complete orthogonal sets and can be normalized so that both sets may be used as basis sets. Thus, the two sets are related by a unitary transformation and we may write

$$|j, m, j_1, j_2\rangle = \sum_{m_1, m_2} \langle j_1 j_2 m_1 m_2 | jm\rangle |j_1, m_1, j_2, m_2\rangle \ . \tag{17.4.23}$$

The coefficients $\langle j_1 j_2 m_1 m_2 | jm\rangle$ are the matrix elements of the unitary transformation connecting the two basis sets and are referred to variously as "Clebsch-Gordon" ," Wigner" or "Vector-addition" coefficients. The inverse of equation (17.4.23) is written as:

$$|j_1, m_1, j_2, m_2\rangle = \sum_{jm} \langle jm | j_1 j_2 m_1 m_2\rangle |j, m, j_1, j_2\rangle \ . \tag{17.4.24}$$

Thus, we immediately have that

$$\sum_{jm} \langle j_1 j_2 m_1 m_2 | jm\rangle \langle jm | j_1 j_2 m_1' m_2'\rangle = \delta_{m_1 m_1'} \delta_{m_2 m_2'} \tag{17.4.25}$$

and

$$\sum_{m_1, m_2} \langle jm | j_1 j_2 m_1 m_2\rangle \langle j_1 j_2 m_1 m_2 | j' m'\rangle = \delta_{jj'} \delta_{mm'} \ . \tag{17.4.26}$$

Clearly the range of summation in (17.4.23) and (17.4.26) is over $-j_1 \leq m_1 \leq j_1$ and $-j_2 \leq m_2 \leq j_2$ since that is the full range of m_1 and m_2. On the other hand we have yet to find the range for j and m. We now do this by a counting procedure as well as by explicitly constructing some of the vectors $|j, m, j_1, j_2\rangle$ in terms of the vectors $|j_1, m_1, j_2, m_2\rangle$. We first notice that the vector

$$|j_1, m_1, j_2, m_2\rangle = |j_1, m_1\rangle |j_2, m_2\rangle \tag{17.4.27}$$

where $|j_i, m_i\rangle$ are the eigenkets of J_i^2 and J_{iz}. This is easily seen to be the case since J_1^2, J_{1z} operate only on $|j_1, m_1\rangle$ and J_2^2, J_{2z} operate only on $|j_2, m_2\rangle$. Furthermore, each of these vectors is also an eigenvector of J_z with eigenvalue $(m_1 + m_2)\hbar$ since

$$J_z = J_{1z} + J_{2z} \tag{17.4.28}$$

and hence

$$\begin{aligned} J_z |j_1, m_1, j_2, m_2\rangle &= J_{1z} |j_1, m_1\rangle |j_2, m_2\rangle + J_{2z} |j_1, m_1\rangle |j_2, m_2\rangle \\ &= (m_1 + m_2)\hbar |j_1, m_1\rangle |j_2, m_2\rangle \ . \end{aligned} \tag{17.4.29}$$

The maximum values of m_1,m_2 are j_1 and j_2 and therefore the maximum eigenvalue m of J_z is

$$m_{\max} = j_1 + j_2 \ . \tag{17.4.30}$$

But, since **J** satisfies the standard algebra of an angular momentum operator it follows from this that the maximum value of j (corresponding to the eigenvalue of J^2) is also $j_1 + j_2$. Hence,

$$j_{\max} = j_1 + j_2 \ . \tag{17.4.31}$$

We now show that in fact j assumes once and only once all the values

$$j = j_1 + j_2 \, , j_1 + j_2 - 1 \, , j_1 + j_2 - 2 \, , \ldots , |j_1 - j_2| \, . \tag{17.4.32}$$

It is clear that by repeatedly applying the operator J_- to the state

$$\begin{aligned} |j_{\max} m_{\max} j_1 j_2\rangle &= |j_1 + j_2, j_1 + j_2, j_1 j_2\rangle \\ &= |j_1, j_1; j_2, j_2\rangle \\ &= |j_1, j_1\rangle |j_2, j_2\rangle \end{aligned} \tag{17.4.33}$$

we get all states $|j_1 + j_2, m, j_1 j_2\rangle$ with m running from $(j_1 + j_2)$ to $-(j_1 + j_2)$. Thus, there are $2j + 1$ different possible values of m for $j = j_1 + j_2$. The same argument holds for any other j value so that there are $2j + 1$ different values of m for every state of definite j.

Now consider states with $m = j_1 + j_2 - 1$. These can be obtained in two ways; namely from $m_1 = j_1$, $m_2 = j_2 - 1$ or $m_1 = j_1 - 1, m_2 = j_2$. Hence, there must be two and only two states of definite j with these m values. They are the states with

$$\begin{array}{ll} j = j_1 + j_2 & m = j - 1 \\ j = j_1 + j_2 - 1 & m = j \ . \end{array} \tag{17.4.34}$$

Similarly for the case of $m = j_1 + j_2 - 2$ there are exactly three states of definite m_1, m_2 values corresponding to the pairs $(j_1 \, , j_2 - 2)$, $(j_1 - 1 \, , j_2 - 1)$, $(j_1 - 2 \, , j_2)$. Thus, there are also exactly three states of definite j, namely

$$\begin{array}{ll} j = j_1 + j_2 & m = j - 2 \\ j = j_1 + j_2 - 1 & m = j - 1 \\ j = j_1 + j_2 - 2 & m = j \ . \end{array} \tag{17.4.35}$$

Similarly for $m = j_1 + j_2 - 3$ there correspond exactly four states of definite m_1, m_2 values, namely, $(j_1 \, , j_2 - 3)$, $(j_1 - 1 \, , j_2 - 2)$, $(j_1 - 2 \, , j_2 - 1)$, $(j_1 - 3 \, , j_2)$ and hence also exactly four states of definite j, namely

$$\begin{array}{ll} j = j_1 + j_2 & m = j - 3 \\ j = j_1 + j_2 - 1 & m = j - 2 \\ j = j_1 + j_2 - 2 & m = j - 1 \\ j = j_1 + j_2 - 3 & m = j \ . \end{array} \tag{17.4.36}$$

Continuing in this way we finally arrive at states with $m = j_1 - j_2$ where for convenience we assumed $j_1 \geq j_2$. In this case there are exactly $2j_2 + 1$ states of definite m_1, m_2 values. They correspond to the pairs $(m_1, m_2) = (j_1 \, , -j_2)$, $(j_1 - 1 \, , -j_2 + 1)$, $\ldots$ $(j_1 - j_2 + 1 \, , 1)$, $(j_1 - j_2 - 2 \, , 2)$, $\ldots$, $(j_1 - 2j_2 + 1 \, , j_2 - 1)$, $\ldots$,

$(j_1 - 2j_2\,, j_2)$. It therefore follows as before that there are also exactly $2j_2+1$ states of definite j, namely

$$\begin{array}{ll} j = j_1 + j_2 & m = j_1 - 2j_2 \\ j = j_1 + j_2 - 1 & m = j_1 - 2j_2 + 1 \\ \dots & \dots \\ j = j_1 + j_2 - 2j_2 & m = j_1 - 2j_2 + j_2\,. \end{array} \tag{17.4.37}$$

Thus, the possible values of j are given in (17.4.32). As an additional check we can count the number of vectors in the two sets $|j,m,j_1,j_2\rangle$ and $|j_1,m_1,j_2,m_2\rangle$. From the form of the states $|j_1,m_1,j_2,m_2\rangle$ as given in (17.4.33) we see that for fixed (j_1,j_2) there are $(2j_1+1)(2j_2+1)$ states of this type. For the states of type $|j,m,j_1,j_2\rangle$ we have that for fixed (j_1,j_2) there are $(2j+1)$ states for each j value over the whole possible range of j values. Therefore, the number of such states is

$$\begin{aligned} &\sum_{j=|j_1-j_2|}^{j_1+j_2} (2j+1) \\ &= 2\frac{(j_1+j_2)(j_1+j_2+1)}{2} - 2\frac{|j_1-j_2|(|j_1-j_2|-1)}{2} + j_1 + j_2 - (|j_1-j_2|-1) \\ &= (2j_1+1)(2j_2+1) \end{aligned} \tag{17.4.38}$$

as desired.

With this result we have arrived at the fundamental addititian theorem for angular momenta.

Fundamental Addition Theorem for Angular Momenta.

In the $(2j_1+1)(2j_2+1)$-dimensional space spanned by the basis set $|j,m,j_1,j_2\rangle$ with j_1 and j_2 fixed, the possible values of j are

$$j = j_1 + j_2\,, j_1 + j_2 - 1\,, j_1 + j_2 - 2\,, \dots, |j_1 - j_2|\,.$$

Corresponding to each of these j values there are $2j+1$ vectors obtained by repeatedly applying J_- to the state $|jjj_1j_2\rangle$.

17.5 Two Spin $\frac{1}{2}$ States

In this case we have $j_1 = j_2 = \frac{1}{2}$. The four possible states in the $|j_1,m_1;j_2,m_2\rangle$ representation are:

$$|\frac{1}{2},\frac{1}{2}\rangle|\frac{1}{2},\frac{1}{2}\rangle\ ,\ |\frac{1}{2},\frac{1}{2}\rangle|\frac{1}{2},-\frac{1}{2}\rangle\ ,\ |\frac{1}{2},-\frac{1}{2}\rangle|\frac{1}{2},\frac{1}{2}\rangle\ ,\ |\frac{1}{2},-\frac{1}{2}\rangle|\frac{1}{2},-\frac{1}{2}\rangle\ .$$

In the $|j,m;j_1,j_2\rangle$ representation the four possible states are:

$$|1,1;\frac{1}{2},\frac{1}{2}\rangle\ ,\ |1,0;\frac{1}{2},\frac{1}{2}\rangle\ ,\ |1,-1;\frac{1}{2},\frac{1}{2}\rangle\ ,\ |0,0;\frac{1}{2},\frac{1}{2}\rangle\ .$$

The state of highest weight is given by

$$|1,1;\frac{1}{2},\frac{1}{2}\rangle = |\frac{1}{2},\frac{1}{2}\rangle|\frac{1}{2},\frac{1}{2}\rangle \ . \tag{17.5.39}$$

The other two states with total $S = 1$ are obtained by applying $S_- = S_{1-} + S_{2-}$ to both sides of equation (17.5.39). This yields

$$\begin{aligned} &\sqrt{1(1+1) - 1(1-1)}\hbar|1,0;1\frac{1}{2},\frac{1}{2}\rangle \\ = \ &\sqrt{\frac{1}{2}(\frac{1}{2}+1) - \frac{1}{2}(\frac{1}{2}-1)}\left\{|\frac{1}{2},-\frac{1}{2}\rangle|\frac{1}{2},\frac{1}{2}\rangle + |\frac{1}{2},\frac{1}{2}\rangle|\frac{1}{2},-\frac{1}{2}\rangle\right\} \end{aligned} \tag{17.5.40}$$

so that

$$|1,0;\frac{1}{2},\frac{1}{2}\rangle = \frac{1}{\sqrt{2}}\left\{|\frac{1}{2},-\frac{1}{2}\rangle|\frac{1}{2},\frac{1}{2}\rangle + |\frac{1}{2},\frac{1}{2}\rangle|\frac{1}{2},-\frac{1}{2}\rangle\right\} \ . \tag{17.5.41}$$

Similarly applying S_- once more to (17.5.41) or else realizing that both spins 1/2 must point down to get a total z-component of spin (i.e. S_z) of $-1\hbar$ we find that

$$|1,-1;\frac{1}{2},\frac{1}{2}\rangle = |\frac{1}{2},-\frac{1}{2}\rangle|\frac{1}{2},-\frac{1}{2}\rangle \ . \tag{17.5.42}$$

The singlet ($S = 0$) state must be orthogonal to the above three states. A simple calculation then shows that

$$|0,0;\frac{1}{2},\frac{1}{2}\rangle = \frac{1}{\sqrt{2}}\left\{|\frac{1}{2},-\frac{1}{2}\rangle|\frac{1}{2},\frac{1}{2}\rangle - |\frac{1}{2},\frac{1}{2}\rangle|\frac{1}{2},-\frac{1}{2}\rangle\right\} \ . \tag{17.5.43}$$

In writing equation (17.5.43) we have made an arbitrary choice of phase.

To display the matrix of Clebsch-Gordon coefficients we rewrite equations (17.5.39) to (17.5.43) as one matrix equation. Thus,

$$\begin{pmatrix} |1,1;1/2,1/2\rangle \\ |1,0;1/2,1/2\rangle \\ |1,-1;1/2,1/2\rangle \\ |0,0;1/2,1/2\rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} |1/2,1/2\rangle|1/2,1/2\rangle \\ |1/2,-1/2\rangle|1/2,1/2\rangle \\ |1/2,1/2\rangle|1/2,-1/2\rangle \\ |1/2,-1/2\rangle|1/2,-1/2\rangle \end{pmatrix} . \tag{17.5.44}$$

Clearly the 4×4 matrix $\langle 1/2, 1/2, m_1, m_2|jm\rangle$ is unitary .

17.6 Spin $\frac{1}{2}$ + Orbital Angular Momentum

The only examples of interest to us will be the coupling of orbital and spin angular momentum for an electron. Since $S = 1/2$ the only possible j values are $l \pm 1/2$, for $l \neq 0$. If $l = 0$ only $j = 1/2$ is possible. We shall simplify the notation somewhat in this case and denote the states $|l \pm 1/2, m; l, 1/2\rangle$ simply by $|l \pm 1/2, m; l\rangle$. Thus, again the state of maximum weight (maximum J_z) is given by

$$|l+\frac{1}{2}, l+\frac{1}{2}; l\rangle = |l,l\rangle|\frac{1}{2},\frac{1}{2}\rangle \, . \tag{17.6.45}$$

Now using the matrix representation

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

for $|\frac{1}{2}, \frac{1}{2}\rangle$ and the coordinate representation $Y_{l,l}$ for $|l, l\rangle$ we have the representation

$$\begin{pmatrix} Y_{l,l} \\ 0 \end{pmatrix}$$

for the state $|l + \frac{1}{2}, l + \frac{1}{2}; l\rangle$. All other states $|l + \frac{1}{2}, m; l\rangle$ are now obtained from this state by repeated application of

$$J_- = L_- + S_- \quad . \tag{17.6.46}$$

Construction of the state $|l - \frac{1}{2}, l - \frac{1}{2}; l\rangle$, from which all states $|l - \frac{1}{2}, m; l\rangle$ are again obtainable by applying J_-, is somewhat more difficult. (See problem 17.1.) The answer, in fact, is

$$|l - \frac{1}{2}, l - \frac{1}{2}; l\rangle = \frac{1}{\sqrt{2l+1}}|l, l-1\rangle|\frac{1}{2}, \frac{1}{2}\rangle - \sqrt{\frac{2l}{2l+1}}|l, l\rangle|\frac{1}{2}, -\frac{1}{2}\rangle \, . \tag{17.6.47}$$

17.7 The Weak-Field Zeeman Effect

We are now finally able to consider the Hamiltonian (17.2.14) which reads:

$$\begin{aligned} H &= \frac{\mathbf{p}^2}{2m} + V(r) + f(r)\mathbf{S} \cdot \mathbf{L} + \frac{e}{2mc}\mathbf{B} \cdot (\mathbf{L} + 2\mathbf{S}) \\ &= \frac{\mathbf{p}^2}{2m} + V(r) + f(r)\mathbf{S} \cdot \mathbf{L} + \frac{e}{2mc}\mathbf{B} \cdot (\mathbf{J} + \mathbf{S}) \end{aligned} \tag{17.7.48}$$

where

$$f(r) = \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr} \tag{17.7.49}$$

is the radial part of the spin-orbit term.

We choose the z-direction to be along the direction of the **B** field which is assumed constant and uniform. Then, defining the Larmor frequency

$$\omega = \frac{eB}{2mc} \tag{17.7.50}$$

the last term becomes $\omega(J_z + S_z)$. The eigenstates of the Hamiltonian without the last term may be labelled $|n, l \pm 1/2, m, l\rangle$ where n is the principal quantum number and j is $l \pm 1/2$. If we treat the term due to the magnetic field, namely $\omega(J_z + S_z)$ by means of first order perturbation theory we get an energy shift ΔE given by

$$\Delta E = \omega\langle n, l \pm 1/2, m, l|J_z + S_z|n, l \pm 1/2, m, l\rangle \tag{17.7.51}$$

or

$$\Delta E = \hbar\omega m + \omega\langle S_z\rangle \ . \qquad (17.7.52)$$

To evaluate this further we must calculate explicitly the spin dependence of the states. Thus, we need a formula of the form

$$|n, l \pm 1/2, m, l\rangle = \sum_{m_l, m_s} \langle l, 1/2, m_l, m_s | l \pm 1/2, m\rangle |n, l, m_l, m_s\rangle \ . \qquad (17.7.53)$$

Rather than evaluate the Clebsch-Gordon coefficients $\langle l, 1/2, m_l, m_s | l \pm 1/2, m\rangle$ we calculate these states directly. As a first step we write the operators corresponding to the total angular momentum $\vec{J} = \vec{L} + \vec{S}$ in matrix form.

$$J_z = \begin{pmatrix} L_z + \hbar/2 & 0 \\ 0 & L_z - \hbar/2 \end{pmatrix} \ . \qquad (17.7.54)$$

To go along with this we write the state $|n, l \pm 1/2, m, l\rangle = |n, j, m, l\rangle$ in configuration space as

$$\psi_{njml} = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \ . \qquad (17.7.55)$$

Then, using

$$\begin{aligned} (L_z + \hbar/2)\psi_1 &= m\hbar\,\psi_1 \\ (L_z - \hbar/2)\psi_2 &= m\hbar\,\psi_2 \end{aligned} \qquad (17.7.56)$$

we find

$$\begin{aligned} \psi_1 &= R_{nl}(r)\, a_1 Y_{l,m-1/2}(\theta,\varphi) \\ \psi_2 &= R_{nl}(r)\, a_2 Y_{l,m+1/2}(\theta,\varphi) \ . \end{aligned} \qquad (17.7.57)$$

We now fix the constants a_1 and a_2 by requiring that ψ_{njml} should also be an eigenfunction of

$$\vec{J}^2 = (\vec{L} + \vec{S})^2 = \begin{pmatrix} \vec{L}^2 + \frac{3}{4}\hbar + \hbar L_z & \hbar(L_x - iL_y) \\ \hbar(L_x + iL_y) & \vec{L}^2 + \frac{3}{4}\hbar - \hbar L_z \end{pmatrix} \qquad (17.7.58)$$

with eigenvalue $j(j+1)\hbar^2$. Thus, using the equations

$$L_\pm Y_{l,m} = \sqrt{l(l+1) - m(m \pm 1)}\,\hbar\, Y_{l,m\pm 1} \qquad (17.7.59)$$

we get

$$\begin{aligned} [j(j+1) - l(l+1) - m - 1/4]\, a_1 - \sqrt{(l+1/2)^2 - m^2}\, a_2 &= 0 \\ \sqrt{(l+1/2)^2 - m^2}\, a_1 - [j(j+1) - l(l+1) + m - 1/4]\, a_2 &= 0 \ . \end{aligned} \qquad (17.7.60)$$

For a solution to exist we need that the determinant vanishes. This yields that

$$j = l + 1/2 \quad \text{or} \quad j = l - 1/2 \ .$$

For $j = l + 1/2$ we get, up to a normalization factor,

$$\begin{aligned} a_1 &= \sqrt{l+m+1/2} \\ a_2 &= \sqrt{l-m+1/2}\ . \end{aligned} \tag{17.7.61}$$

So,

$$\psi_{n,j=l+1/2,m,l} = \frac{R_{nl}(r)}{\sqrt{2l+1}} \begin{pmatrix} \sqrt{l+m+1/2}\, Y_{l,m-1/2} \\ \sqrt{l-m+1/2}\, Y_{l,m+1/2} \end{pmatrix} . \tag{17.7.62}$$

Here we have included the normalization factor $1/\sqrt{2l+1}$. Similarly for $j = l - 1/2$ we get

$$\psi_{n,j=l-1/2,m,l} = \frac{R_{nl}(r)}{\sqrt{2l+1}} \begin{pmatrix} \sqrt{l-m+1/2}\, Y_{l,m-1/2} \\ -\sqrt{l+m+1/2}\, Y_{l,m+1/2} \end{pmatrix} . \tag{17.7.63}$$

We can now complete our calculation of ΔE. Thus, for example, for the state $|n, j = l + 1/2, m, l\rangle$ we have

$$\begin{aligned} &\langle S_z \rangle = \\ &\frac{1}{2l+1}\left(\begin{pmatrix} \sqrt{l+m+\frac{1}{2}}\, Y_{l,m-\frac{1}{2}}(\theta,\varphi) \\ \sqrt{l-m+\frac{1}{2}}\, Y_{l,m+\frac{1}{2}}(\theta,\varphi) \end{pmatrix}, S_z \begin{pmatrix} \sqrt{l+m+\frac{1}{2}}\, Y_{l,m-\frac{1}{2}}(\theta,\varphi) \\ \sqrt{l-m+\frac{1}{2}}\, Y_{l,m+\frac{1}{2}}(\theta,\varphi) \end{pmatrix} \right) \end{aligned} \tag{17.7.64}$$

so that in this case

$$\begin{aligned} &\langle S_z \rangle = \\ &\frac{\hbar}{2}\frac{1}{2l+1}\left[(l+m+\frac{1}{2})(Y_{l,m-\frac{1}{2}}, Y_{l,m-\frac{1}{2}}) - (l-m+\frac{1}{2})(Y_{l,+\frac{1}{2}}, Y_{l,+\frac{1}{2}}) \right] \\ &\quad = \frac{m\hbar}{2l+1} \end{aligned} \tag{17.7.65}$$

and

$$\Delta E_{l+1/2,m,l} = \hbar\omega m \left(1 + \frac{1}{2l+1} \right) . \tag{17.7.66}$$

Similarly for the state $|n, j = l - 1/2, m, l\rangle$

$$\begin{aligned} &\langle S_z \rangle = \\ &\frac{1}{2l+1}\left(\begin{pmatrix} \sqrt{l-m+\frac{1}{2}}\, Y_{l,m-\frac{1}{2}}(\theta,\varphi) \\ -\sqrt{l+m+\frac{1}{2}}\, Y_{l,m+\frac{1}{2}}(\theta,\varphi) \end{pmatrix}, S_z \begin{pmatrix} \sqrt{l-m+\frac{1}{2}}\, Y_{l,m-\frac{1}{2}}(\theta,\varphi) \\ -\sqrt{l+m+\frac{1}{2}}\, Y_{l,m+\frac{1}{2}}(\theta,\varphi) \end{pmatrix} \right) \end{aligned} \tag{17.7.67}$$

so

$$\begin{aligned} &\langle S_z \rangle = \\ &\frac{\hbar}{2}\frac{1}{2l+1}\left[(l-m+\frac{1}{2})(Y_{l,m-\frac{1}{2}}, Y_{l,m-\frac{1}{2}}) - (l+m+\frac{1}{2})(Y_{l,+\frac{1}{2}}, Y_{l,+\frac{1}{2}}) \right] \\ &\quad = -\frac{m\hbar}{2l+1} \end{aligned} \tag{17.7.68}$$

and in this case we obtain for the weak field Zeeman Effect

$$\Delta E_{l-1/2,m,l} = \hbar\omega m \left(1 - \frac{1}{2l+1}\right) \quad . \tag{17.7.69}$$

17.8 The Aharonov-Bohm Effect

In Chapter 16 as well as in this chapter we considered the interaction of a charged particle with an electromagnetic field. We now again look at this interaction in some detail. The Aharonov-Bohm (AB) effect was first mentioned by Ehrenberg and Siday [17.4], but was really brought to the foront by the work of Aharonov and Bohm [17.5]. A controversy over this effect raged for more than two decades but has been settled, by some fine experimental work [17.6], in favour of the conventional quantum mechanical interpretation.

In the AB effect a charged particle, such as an electron, is seen to be affected by electromagnetic potentials in regions in which no electromagnetic fields exist. An important point here is that for this effect to occur, space no longer forms a simply connected region. The effect is crucially dependent on the non-trivial topology of space.

A second important point, which we now recall, is that *the wave-function for angular momenta must be single-valued.* This was discussed in chapter 9 (see reference [9.1]). We now turn to two classes of explicit problems that illustrate this effect. In this discussion we follow closely the original papers of Aharonov and Bohm.

Consider a charged particle, such as an electron, surrounded by a conductor which is connected to a time-varying potential. Inside the conductor the electric potential can only vary in time, not in space. Thus, the Hamiltonian of the charged particle, inside the conductor, has an additional term $eV(t)$ so that

$$H = H_0 + eV(t) \tag{17.8.70}$$

where H_0 is the Hamiltonian when the conductor is grounded. If the wave-function corresponding to H_0 is $\psi_0(x,t)$ then the solution $\psi(x,t)$, for the full Hamiltonian H, is given by

$$\psi(x,t) = \psi_0(x,t)\exp(-\frac{ie}{\hbar}S) \quad , \quad S = \int V(t)\,dt \quad . \tag{17.8.71}$$

The addition of the phase factor $\exp(-ie/\hbar\, S)$ has no observable physical effect. Now consider the case of an electron beam that is split so as to pass through two different regions (fig. 17.1) each of which is partly surrounded by a conductor. In this case, space is no longer simply connected, there is a "hole" in space, a region inaccessible to the electrons. Now apply a different time-varying potential to each of the conductors and only during the time that the electrons are completely inside the conductors. The idea is to avoid exerting any forces on the electrons. Let the solution, when both conductors are grounded, be

$$\psi(x,t) = \psi_{1,0}(x,t) + \psi_{2,0}(x,t) \tag{17.8.72}$$

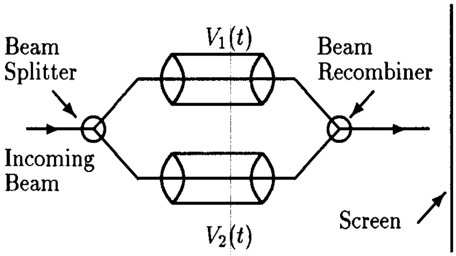

Figure 17.1: The electric Aharonov-Bohm effect.

where $\psi_{1,0}(x,t)$ and $\psi_{2,0}(x,t)$ represent those parts of the wavefunctions that correspond to the paths 1 and 2 respectively. The solutions with potentials turned on, as described above, are then

$$\psi(x,t) = \psi_{1,0}(x,t)\exp(-i(e/\hbar)\,S_1) + \psi_{2,0}(x,t)\exp(-i(e/\hbar)\,S_2) \quad (17.8.73)$$

where

$$S_1 = \int V_1(t)\,dt \quad , \quad S_2 = \int V_2(t)\,dt \,. \quad (17.8.74)$$

This time the phases, from the two different regions, are different and thus produce interference when the two beams are recombined. The phase difference is given by

$$\frac{e}{\hbar}(S_1 - S_2) = \oint V(t)\,dt \,. \quad (17.8.75)$$

A similar result holds in the case of purely magnetic fields and only a vector potential, as can be seen from the relativistic generalization of the integral in (17.8.75), namely

$$\frac{e}{\hbar}\oint\left[V(t)\,dt - \frac{1}{c}\mathbf{A}\cdot d\mathbf{x}\right] \,. \quad (17.8.76)$$

Thus, in the case of only a magnetic field, the interesting quantity is

$$\frac{e}{\hbar}S = -\frac{e}{\hbar c}\oint \mathbf{A}\cdot d\mathbf{x} = -\frac{e}{\hbar c}\int \mathbf{B}\cdot d\mathbf{a} = -\frac{e}{\hbar c}\Phi \quad (17.8.77)$$

where Φ is the total magnetic flux passing through the area enclosed by the loop in the loop integral.

To see how this comes about we again consider a coherent beam of electrons which is split into two parts as in figure 17.2 so that each part passes on either side of a solenoid containing a quantity of magnetic flux Φ before being recombined. This time the Hamiltonian of interest is

$$H = \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 \,. \quad (17.8.78)$$

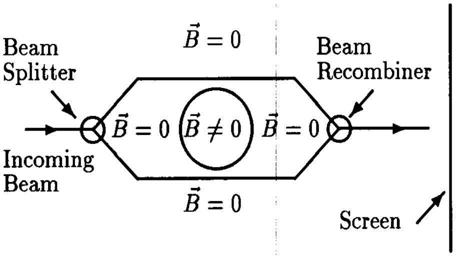

Figure 17.2: The magnetic Aharonov-Bohm effect.

If the magnetic field is a pure gauge field,

$$\mathbf{B} = \nabla \times \mathbf{A} = 0 \tag{17.8.79}$$

(so that also $\Phi = 0$) and if the region is simply connected then the wavefunction ψ for the above Hamiltonian can always be found from

$$\psi = \psi_0 \exp(-i(e/\hbar)\, S) \tag{17.8.80}$$

where

$$\nabla S = \frac{1}{c}\mathbf{A} \tag{17.8.81}$$

and ψ_0 is the wavefunction corresponding to $\mathbf{A} = 0$. On the other hand, for the situation of interest, where space is divided into two regions such that the magnetic field $\mathbf{B}$ is nonzero in an infinitely long tube and zero outside this tube, the solution (17.8.80) above would be a multiple-valued function, for any path that circles the solenoid, and therefore not an acceptable solution. However, if we consider the beam of electrons passing on either side of this solenoid then each part of the beam is in a simply connected region and therefore single-valued. Thus, each portion of the beam is of the form given by (17.8.80), and the recombined beam is of the form

$$\psi = \psi_{1,0} \exp(-i(e/\hbar)\, S_1) + \psi_{2,0} \exp(-i(e/\hbar)\, S_2) \tag{17.8.82}$$

where S_1 and S_2 are given by

$$\frac{1}{c}\int \mathbf{A}\cdot d\mathbf{x} \tag{17.8.83}$$

along their respective paths. The interference between the recombined beams is therefore determined by the phase difference

$$\frac{e}{\hbar}(S_1 - S_2) = \frac{e}{\hbar c}\oint \mathbf{A}\cdot d\mathbf{x} = \frac{e}{\hbar c}\Phi \ . \tag{17.8.84}$$

It is important to note here, that in spite of the appearance of potentials in the expression (17.8.75) and (17.8.84) for the phase difference, the results are nevertheless gauge invariant and depend only on the total magnetic flux Φ.

This result is the magnetic AB effect. The important point here is that the theory predicts that a charged particle, such as an electron, can be influenced by potentials even if the particle only passes through regions where the electromagnetic field is zero. In a multiply connected region of space, free of electromagnetic fields, a charged particle will still experience measurable effects that depend on the potentials. The results are, of course, still gauge invariant and depend only on the total flux $\Phi = \oint \mathbf{A} \cdot d\mathbf{x}$. However, the particle never experiences the electromagnetic fields or the forces exerted by them. The conclusion to be reached is that it is not only the electromagnetic fields that are physically measurable but also gauge invariant quantities such as the magnetic flux.

17.9 Problems

17.1 Show that the operator

$$T = A[j_2 J_{1-} - j_1 J_{2-}]$$

where A is a normalization constant, has the property that

$$T|j, j, j_1, j_2\rangle = |j-1, j-1, j_1, j_2\rangle$$

for $j = j_1 + j_2$.

17.2 Evaluate the expectation value $\langle S_z \rangle$ for the states $|l \pm 1/2, m, l\rangle$ and hence the shift in energy due to a uniform magnetostatic field.

17.3 A particle of angular momentum 1/2 is coupled to a particle of angular momentum 1. List the states that are eigenstates of

$$J^2 = (\mathbf{J}_1 + \mathbf{J}_2)^2$$

and

$$J_z = J_{1z} + J_{2z}$$

and express them in terms of the eigenstates of

$$(J_1^2, J_{1z}) \quad \text{and} \quad (J_2^2, J_{2z}) .$$

17.4 Consider a set of three operators T_m $m = 1, 0, -1$ such that

$$T_m^\dagger = T_m$$

$$[J_\pm, T_m] = \sqrt{2 - m(m \pm 1)}\, \hbar T_m$$

$$[J_z, T_m] = m\hbar\, T_m$$

where **J** are the total angular momentum operators. Evaluate the total m' , m'' dependence of the matrix elements $\langle j, m'|T_m|j, m''\rangle$.
Hint: Express T_m in terms of 3×3 matrices. This is an example of the Wigner-Eckart Theorem.

17.5 Consider the unitary operator

$$R_n(\varphi) = e^{i(\mathbf{J}\cdot\mathbf{n}\varphi/\hbar)}$$

where **J** is the angular momentum operator.
a) If $j = 1/2$ expand $R_n(\varphi)$ in a Taylor series to obtain a simpler expression and apply it to the states

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

What is the effect of $R_n(\varphi)$?
b) If $j = 1$ repeat part a) but consider the states

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} , \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} .$$

17.6 A particle of total angular momentum $j_1 = 1/2$ is coupled to another particle with total angular momentum $j_2 = 3/2$. What are the states of possible total j? Express all the states with the lowest possible j in terms of the states $|j_1, m_1\rangle$, $|j_2, m_2\rangle$.

17.7 An electron (spin = 1/2) is in a state of either $l = 0$ or $l = 1$. Express all states of total angular momentum $|j, m_j\rangle$ in terms of the states $|l, m_l\rangle|1/2, m_s\rangle$ where $l = 0$ or $l = 1$.

17.8 Show that if

$$\mathbf{A} = -\frac{1}{2}\mathbf{r} \times \mathbf{B}$$

where **B** is a constant vector, then

$$\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p} = \mathbf{B} \cdot \mathbf{L} .$$

17.9 Use induction to show that

$$|n, j, m, l\rangle = \sqrt{\frac{(j+m)!}{(2j)!(j-m)!}}\,\hbar^{m-j}\, J_-^{j-m}|n, j, j, l\rangle \ .$$

17.10 Show that if **A** is a vector operator such that

$$[J_x, A_x] = 0 \quad , \quad [J_x, A_y] = i\hbar A_z \quad , \quad [J_x, A_y] = -i\hbar A_y$$

and cyclic permutations. Then,

a)

$$[J^2, [J^2, \mathbf{A}]] = 2\hbar^2(J^2\mathbf{A} + \mathbf{A}J^2) - 4\hbar^2(\mathbf{A}\cdot\mathbf{J})\mathbf{J}$$

b) Use this result to show that

$$\langle JM|S_z|JM\rangle = M\hbar\frac{\langle JM|\mathbf{S}\cdot\mathbf{J}|JM\rangle}{J(J+1)\hbar^2}$$

and

c) hence evaluate the matrix element for the weak-field Zeeman effect

$$\Delta E = \frac{\mu_B B}{\hbar}\{M\hbar + \langle nLSJM|S_z|nLSJM\rangle\} = \mu_B BMg$$

where

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}$$

is the *Landé g-factor.*

Bibliography

[17.1] G.E. Uehlenbeck and S. Goudsmit, Naturwiss. **13**, 953 (1925); ibid Nature **117**, 264 (1926).
The Thomas precession was presented in

[17.2] L.H. Thomas, Nature **107**, 514 (1926).

[17.3] A detailed treatment of atomic spectra is to be found in the book by Condon and Shortley [15.1].
The addition of angular momenta is simplest when viewed from the point of view of group theory. A representative text for physicists is
M. Hammermesh, *Group Theory and its Applications to Physical Problems* - Addison-Wesley Publishing Co., Inc., Reading, Mass., U.S.A. (1962).
In using tables of Clebsch-Gordon coefficients great care must be exercised since several different phase-conventions are in use. Besides the one used in Condon and Shortley common ones are also found in Wigner's book d in Chapter 12. [12.2] as well as in:
J.M. Blatt and V.F. Weisskopf, *Theoretical Nuclear Physics* -Wiley, New York (1952).
In practice it is very common to use standard computer codes that generate all the desired Clebsch-Gordon coefficients.

[17.4] W. Ehrenberg and R.E. Siday, Proc. Phys. Soc. London, **B62**, 8 (1949).

[17.5] Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1959).
ibid, **123**, 1511 (1961).
ibid, **125**, 2192 (1962).
ibid, **130**, 1625 (1963)

[17.6] The experimental work as well as numerous references are well-described in M.Peshkin and A.Tonomura, *The Aharonov-Bohm Effect*, Springer-Verlag (1989).

Chapter 18

Scattering Theory - Time Dependent

18.1 Introduction

So far, except for a few examples in Chapters 4 and 5, we have been concerned almost exclusively with the discrete part of the energy spectrum. In this chapter we commence a discussion of the continuous spectrum. Thus, we do not solve for the energy; in fact, we solve for the wave-function corresponding to certain initial conditions. These initial conditions correspond to a current of particles incident on some potential and a current of particles scattered by the same potential.

A very large number of microscopic phenomena have their origin in the collisions of particles. For example such diverse properties as the conductivity of metals and the critical masses of nuclear reactions are ultimately determined by scattering phenomena. Furthermore, almost all nuclear and high energy physics experiments are collision experiments and require some form of scattering theory for their interpretation. Thus, scattering theory is one of the most important tools of a modern physicist.

The time-dependent formulation allows a very intuitive approach since the concepts used correspond closely to those used in classical scattering theory. However, for computational purposes, the time-independent formulation is more convenient. Therefore those readers more interested in the applications of scattering theory may omit all of this chapter except section 18.7 and proceed directly to Chapter 19. The only other result used later is equation (18.9.111). This section may be then read with profit after section 3 of Chapter 19.

To bring out the similarity between the time-dependent formulation of scattering theory and classical scattering theory we begin with a quick review of the main physical concepts used in the classical scattering problem.

18.2 Classical Scattering Theory

Consider two particles that may interact with each other. Classically the state of this two-particle system at any time t is given by the four vectors $\mathbf{x}_1(t)$, $\mathbf{p}_1(t)$, $\mathbf{x}_2(t)$, $\mathbf{p}_2(t)$ specifying the trajectories of both particles. In practice it is much more convenient to reduce this problem to an equivalent one-body problem using the law of conservation of momentum. Thus, we define

$$\mathbf{X}(t) = \frac{m_1\mathbf{x}_1(t) + m_2\mathbf{x}_2(t)}{m_1 + m_2} \tag{18.2.1}$$

$$\mathbf{x}(t) = \mathbf{x}_1(t) - \mathbf{x}_2(t) \tag{18.2.2}$$

$$\mathbf{P}(t) = \mathbf{p}_1(t) + \mathbf{p}_2(t) \tag{18.2.3}$$

$$\mathbf{p}(t) = \frac{m_2}{m_1 + m_2}\mathbf{p}_1(t) - \frac{m_1}{m_1 + m_2}\mathbf{p}_2(t) \ . \tag{18.2.4}$$

Conservation of momentum then implies that $\mathbf{P}(t)$ is a constant independent of time, say $\mathbf{P}_{in}$, and hence $\mathbf{X}(t)$ depends linearly on t so that

$$\mathbf{X}(t) = \mathbf{X}_{in} + \frac{\mathbf{P}_{in}}{m_1 + m_2} t \ . \tag{18.2.5}$$

Equation (18.2.5) is simply a statement of the fact that the centre of mass moves like a free particle with momentum $\mathbf{P}_{in}$, mass $M = m_1 + m_2$ and position $\mathbf{X}_{in}$ at $t = 0$. One now has to solve for the trajectory of the equivalent one-body problem described by the variables $\mathbf{x}(t)$, $\mathbf{p}(t)$.

In a scattering problem involving an interaction of short range the trajectory will have the qualitative features shown in fig. 18.1. Far from the interaction region the fictitious particle corresponding to the *reduced mass*

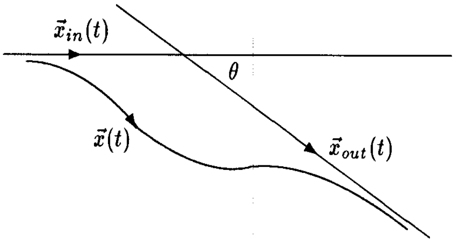

Figure 18.1: Classical scattering by a central force.

$$m = \frac{m_1 m_2}{m_1 + m_2}$$

moves like a free particle indicated by the two trajectories $\mathbf{x}_{in}(t)$ and $\mathbf{x}_{out}(t)$. In the interaction region the trajectory is complicated. We can describe the two asymptotic trajectories by

$$\mathbf{x}_{in}(t) = \mathbf{x}_{in}(0) + \frac{\mathbf{p}_{in}}{m} t \tag{18.2.6}$$

$$\mathbf{x}_{out}(t) = \mathbf{x}_{out}(0) + \frac{\mathbf{p}_{out}}{m} t \ . \tag{18.2.7}$$

Due to the assumed short-range nature of the potential we may also impose the asymptotic conditions

$$\mathbf{x}(t) \to \mathbf{x}_{in}(t) \quad \text{for} \quad t \to -\infty \tag{18.2.8}$$

and

$$\mathbf{x}(t) \to \mathbf{x}_{out}(t) \quad \text{for} \quad t \to \infty \ . \tag{18.2.9}$$

This limit has to be defined precisely and is in fact defined by

$$\lim_{t\to-\infty} \| \mathbf{x}(t) - \mathbf{x}_{in}(t) \| = 0 \tag{18.2.10}$$

and

$$\lim_{t\to\infty} \| \mathbf{x}(t) - \mathbf{x}_{out}(t) \| = 0 \tag{18.2.11}$$

where $\| \cdot \|$ here represents the Euclidean norm. From the two asymptotic trajectories we obtain the scattering angle by using

$$\cos\theta = \frac{\mathbf{x}_{in}(t) \cdot \mathbf{x}_{out}(t)}{\| \mathbf{x}_{in}(t) \| \| \mathbf{x}_{out}(t) \|} \ . \tag{18.2.12}$$

This is just the "overlap" between the asymptotic states. We next derive the analogous procedure in quantum mechanics.

18.3 Asymptotic States: Schrödinger Picture

We start immediately with the Hamiltonian for the equivalent one-body problem

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) \ . \tag{18.3.13}$$

As in the classical case we assume that $V(\mathbf{r})$ is short range [1] so that for large $|\mathbf{r}| = r$, $V(\mathbf{r})$ dies out rapidly and H approaches H_0, the Hamiltonian for a free particle

$$H_0 = \frac{\mathbf{p}^2}{2m} \ . \tag{18.3.14}$$

---

[1] A precise definition of a short-range potential is not required, but the condition $r^2 V(r) \to 0$ for $r \to \infty$ is sufficient.

In the classical description the trajectories $\mathbf{x}_{in}(t)$, $\mathbf{x}_{out}(t)$ were in fact determined by H_0. Similarly, in the quantum mechanical case, we assume that there exist states $\Psi_{in}(t)$, $\Psi_{out}(t)$ evolving according to H_0 such that

$$\lim_{t\to-\infty} \| \Psi(t) - \Psi_{in}(t) \| = 0 \tag{18.3.15}$$

$$\lim_{t\to\infty} \| \Psi(t) - \Psi_{out}(t) \| = 0 \ . \tag{18.3.16}$$

These limits, if they exist, define the asymptotic incoming and outgoing states. Here $\Psi(t)$ is a solution of the full Schrödinger equation with the Hamiltonian (18.3.13). The symbol $\| \cdot \|$ here represents the Hilbert space norm and the limits involved are called *strong limits.* (See problem 6.1.) One states this by saying that $\Psi(t)$ converges strongly to $\Psi_{in}(t)$ as $t \to -\infty$ and $\Psi(t)$ converges strongly to $\Psi_{out}(t)$ as $t \to \infty$.

Since H and H_0 are self-adjoint, they determine unitary evolution operators

$$U(t) = e^{-iHt/\hbar} \tag{18.3.17}$$

and

$$U_0(t) = e^{-iH_0t/\hbar} \ . \tag{18.3.18}$$

Thus, the requirements for asymptotic states (equations (18.3.15) and (8.2.29)) may be written

$$\lim_{t\to\pm\infty} \| e^{-iHt/\hbar}\Psi(0) - e^{-iH_0t/\hbar}\Psi_{in,out}(0) \| = 0 \ . \tag{18.3.19}$$

Using the unitarity of these operators we can further write

$$\lim_{t\to\pm\infty} \| e^{iH_0t/\hbar} e^{-iHt/\hbar}\Psi(0) - \Psi_{in,out}(0) \| = 0 \ . \tag{18.3.20}$$

We are thus led in a very natural fashion to consider the two operators

$$\Omega^\dagger_\mp = \lim_{t\to\pm\infty} e^{iH_0t/\hbar} e^{-iHt/\hbar} \ . \tag{18.3.21}$$

These operators are known as the *Møller wave operators.* Just as in the classical case the two asymptotic trajectories gave us the amount of scattering (scattering angle), so also quantum mechanically the amount of scattering is determined by the overlap between the two asymptotic solutions. It is in this respect that the two Møller operators are particularly useful.

18.4 The Møller Wave Operators

We have from the definition of $\Omega_\pm$ that

$$|\Psi_{in}(t)\rangle = \Omega^\dagger_+ |\Psi(t)\rangle \tag{18.4.22}$$

and

$$|\Psi_{out}(t)\rangle = \Omega^\dagger_- |\Psi(t)\rangle \ . \tag{18.4.23}$$

But from (18.3.20) we can also deduce that

$$|\Psi(t)\rangle = \Omega_+|\Psi_{in}(t)\rangle = \Omega_-|\Psi_{out}(t)\rangle \tag{18.4.24}$$

Hence combining these with (18.4.22) and (18.4.23) we get

$$|\Psi_{in}(t)\rangle = \Omega_+^\dagger\Omega_+|\Psi_{in}(t)\rangle \tag{18.4.25}$$

$$|\Psi_{out}(t)\rangle = \Omega_-^\dagger\Omega_-|\Psi_{out}(t)\rangle \ . \tag{18.4.26}$$

We assume of course that the "in" and "out" states span the entire Hilbert space of scattering states so that at any time every scattering state can be written as a linear superposition of states $|\Psi_{in}(t)\rangle$ or $|\Psi_{out}(t)\rangle$. From this and (18.4.25) or (18.4.26) we conclude that

$$\Omega_\pm^\dagger\Omega_\pm = 1 \tag{18.4.27}$$

on the space of scattering states. This does not, however, imply that the operators $\Omega_\pm$ are unitary on the whole Hilbert space since one cannot establish the relation $\Omega_\pm^\dagger\Omega_\pm = 1$ for all physical states even though we do have

$$\Omega_\pm^\dagger\Omega_\pm|\Psi(t)\rangle = |\Psi(t)\rangle \tag{18.4.28}$$

on the space of scattering states $|\Psi(t)\rangle$. The reason for this is, that if H has bound states then the scattering states $|\Psi(t)\rangle$ do not span the entire Hilbert space. All the scattering states are orthogonal to the bound states and these can therefore not be expanded in terms of scattering states. If we now return to the Schrödinger equation for $|\Psi(t)\rangle$ or either $|\Psi_{in}(t)\rangle$ or $|\Psi_{out}(t)\rangle$ we get

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = H|\Psi(t)\rangle \tag{18.4.29}$$

and

$$i\hbar\frac{\partial}{\partial t}|\Psi_{in,out}(t)\rangle = H_0|\Psi_{in,out}(t)\rangle \ . \tag{18.4.30}$$

Substituting (18.4.24) into (18.4.29) and comparing with (18.4.30) we get the so-called *inter-twining property* of the Møller wave operators

$$H\Omega_\pm = \Omega_\pm H_0 \ . \tag{18.4.31}$$

Another very useful relation is obtained by combining (18.4.24) with (18.4.23).

$$|\Psi_{out}(t)\rangle = \Omega_-^\dagger\Omega_+|\Psi_{in}(t)\rangle \ . \tag{18.4.32}$$

This relates the asymptotic states at $t = -\infty$ to those at $t = +\infty$. The combination $\Omega_-^\dagger\Omega_+$ plays an extremely important role in scattering theory and defines the S-operator

$$S = \Omega_-^\dagger\Omega_+ \ . \tag{18.4.33}$$

Evaluated on the *basis of free states* this operator is called the S-matrix.

If instead of combining (18.4.24) with (18.4.23) we had combined (18.4.24) with (18.4.22) we would have obtained instead of (18.4.32) the relation

$$|\Psi_{in}(t)\rangle = \Omega_+^\dagger \Omega_- |\Psi_{out}(t)\rangle \tag{18.4.34}$$

or

$$|\Psi_{in}(t)\rangle = S^\dagger |\Psi_{out}(t)\rangle \ . \tag{18.4.35}$$

Since the sets $\{|\Psi_{in}(t)\rangle\}$ and $\{|\Psi_{out}(t)\rangle\}$ are complete we can conclude that S is unitary

$$SS^\dagger = S^\dagger S = 1 \ . \tag{18.4.36}$$

Using the relation (18.4.31), as well as the fact that H and H_0 are self-adjoint, we find

$$H\Omega_\pm = \Omega_\pm H_0 \quad \text{and} \quad H_0 \Omega_\pm^\dagger = \Omega_\pm^\dagger H \ . \tag{18.4.37}$$

Multiplying the first of these by $\Omega_-^\dagger$ on the left and the second by Ω_+ on the right, we get

$$\Omega_-^\dagger H \Omega_+ = \Omega_-^\dagger \Omega_+ H_0 = SH_0 \tag{18.4.38}$$

and

$$H_0 \Omega_-^\dagger \Omega_+ = H_0 S = \Omega_-^\dagger H \Omega_+ \ . \tag{18.4.39}$$

Thus, we conclude that

$$SH_0 = H_0 S \ . \tag{18.4.40}$$

This means that if we take matrix elements of S with eigenstates of H_0 (basis of free states) the resultant S-matrix elements are independent of time.

18.5 Green's Functions and Propagators

Our Hamiltonian is as before

$$H = H_0 + V(r) \tag{18.5.41}$$

with

$$H_0 = \frac{\mathbf{p}^2}{2m} \ . \tag{18.5.42}$$

We now define four Green's functions or propagators via the equations

$$\left(i\hbar \frac{\partial}{\partial t} - H_0 \right) G_0^\pm(t) = \delta(t)\mathbf{1} \tag{18.5.43}$$

$$\left(i\hbar \frac{\partial}{\partial t} - H \right) G^\pm(t) = \delta(t)\mathbf{1} \tag{18.5.44}$$

and the boundary conditions

$$G_0^+(t) = G^+(t) = 0 \quad \text{for} \quad t < 0 \tag{18.5.45}$$

$$G_0^-(t) = G^-(t) = 0 \quad \text{for} \quad t > 0 \ . \tag{18.5.46}$$

In the usual terminology of differential equations G_0^+, G^+ are *retarded* Green's functions, while G_0^-, G^- are *advanced* Green's functions. They are closely related to the Møller wave operators.

Formal solutions of (18.5.43) and (18.5.44) incorporating the boundary conditions (18.5.45) or (18.5.46) are

$$G_0^+(t) = -\frac{i}{\hbar}\, e^{-iH_0t/\hbar}\, \theta(t) \tag{18.5.47}$$

$$G_0^-(t) = \frac{i}{\hbar}\, e^{-iH_0t/\hbar}\, \theta(-t) \tag{18.5.48}$$

and

$$G^+(t) = -\frac{i}{\hbar}\, e^{-iHt/\hbar}\, \theta(t) \tag{18.5.49}$$

$$G^-(t) = \frac{i}{\hbar}\, e^{-iHt/\hbar}\, \theta(-t) \ . \tag{18.5.50}$$

The function $\theta(t)$ is defined by

$$\theta(t) = \begin{cases} 1 & t > 0 \\ 0 & t < 0 \end{cases} \tag{18.5.51}$$

such that

$$\begin{aligned} \lim_{t\to 0+} \theta(t) &= 1 \\ \lim_{t\to 0-} \theta(t) &= 0 \ . \end{aligned} \tag{18.5.52}$$

These solutions are to be understood in the sense of distributions (see Chapter 9) if matrix elements of these solutions are involved. Clearly they are very closely related to evolution operators and that, in fact, is the reason they are also called *propagators.* Thus, if $|\Psi_0(t)\rangle$ is a solution of the Schrödinger equation

$$\left(i\hbar\frac{\partial}{\partial t} - H_0\right)|\Psi_0(t)\rangle = 0 \tag{18.5.53}$$

then for $t > t'$ we can write

$$|\Psi_0(t)\rangle = i\hbar G_0^+(t-t')|\Psi_0(t')\rangle \ . \tag{18.5.54}$$

Thus G_0^+ propagates states, evolving according to H_0, from one time to a later time. Similarly if $|\Psi(t)\rangle$ is a solution of

$$\left(i\hbar\frac{\partial}{\partial t} - H\right)|\Psi(t)\rangle = 0 \tag{18.5.55}$$

then for $t > t'$

$$|\Psi(t)\rangle = i\hbar G^{+}(t-t')|\Psi(t')\rangle \tag{18.5.56}$$

so that G^+ propagates states, evolving according to H, from one time to a later time.

In a completely analogous fashion we also obtain the relations

$$|\Psi_0(t)\rangle = -i\hbar G_0^{-}(t-t')|\Psi_0(t')\rangle \quad \text{for} \quad t < t' \tag{18.5.57}$$

and

$$|\Psi(t)\rangle = -i\hbar G^{-}(t-t')|\Psi(t')\rangle \quad \text{for} \quad t < t' \ . \tag{18.5.58}$$

Next we show that the Møller operators are appropriate limits of these propagators, corresponding to propagation from a time $-\infty$ to a time t or from t to a time $+\infty$.

Consider a state $|\Psi(t')\rangle$, which evolves with the Hamiltonian H and is therefore a solution of the full Schrödinger equation. We now define the state

$$|\Psi_0(t)\rangle = i\hbar G_0^{+}(t-t')|\Psi(t')\rangle \ . \tag{18.5.59}$$

Then for $t > t'$ $|\Psi_0(t)\rangle$ evolves with the Hamiltonian H_0 and is therefore a solution of the Schrödinger equation with the Hamiltonian H_0. Furthermore, taking the limit $t \to t' + 0$ we find that

$$|\Psi_0(t')\rangle = |\Psi(t')\rangle \tag{18.5.60}$$

since

$$\lim_{t\to 0+} G_0^{+}(t) = -\frac{i}{\hbar} \tag{18.5.61}$$

according to the definition (18.5.52) of the θ-function. Next consider taking the limit $t' \to -\infty$. This limit yields $|\Psi_{in}(t)\rangle$. To see this we use the formal solutions (18.5.50) and (18.5.47) for the Green's function and their properties as propagators given in (18.5.54) and (18.5.55). In particular, using (18.5.58)

$$|\Psi(t')\rangle = -i\hbar G^{-}(t'-t'')|\Psi(t'')\rangle \quad \text{for} \quad t' < t'' \ . \tag{18.5.62}$$

Now we use the definition (18.5.59)

$$|\Psi_0(t)\rangle = i\hbar G_0^{+}(t-t')|\Psi(t')\rangle \quad \text{for} \quad t > t' \tag{18.5.63}$$

and substitute from (18.5.47) and (18.5.50) to obtain

$$\begin{aligned} |\Psi_0(t)\rangle \quad &= \quad e^{-iH_0(t-t')/\hbar}\theta(t-t')e^{-iH(t'-t'')/\hbar}\theta(t'-t'')|\Psi(t'')\rangle \\ &\quad\quad \text{for} \quad t > t'' > t' \ . \end{aligned} \tag{18.5.64}$$

Or since

$$|\Psi_0(t)\rangle = e^{-iH_0(t-t'')/\hbar}|\Psi_0(t'')\rangle \quad \text{for} \quad t > t'' \tag{18.5.65}$$

we obtain

$$|\Psi_0(t'')\rangle = \theta(t-t')\theta(t''-t')e^{iH_0(t'-t'')/\hbar}e^{-iH(t'-t'')/\hbar}|\Psi(t'')\rangle \ . \tag{18.5.66}$$

If we now let $t' \to -\infty$ then the two θ-functions reduce to 1 and we get

$$|\Psi_0(t'')\rangle = \lim_{t'\to-\infty} e^{iH_0(t'-t'')/\hbar}e^{-iH(t'-t'')/\hbar}|\Psi(t'')\rangle \tag{18.5.67}$$

or, by the definition of $\Omega_+^\dagger$

$$|\Psi_0(t'')\rangle = \Omega_+^\dagger|\Psi(t'')\rangle \ . \tag{18.5.68}$$

Thus, dropping the primes, we have

$$|\Psi_0(t)\rangle = |\Psi_{in}(t)\rangle \tag{18.5.69}$$

as required.

In a similar manner one can define a free state $|\Psi_0'(t)\rangle$ by taking the limit

$$|\Psi_0'(t)\rangle = \lim_{t'\to\infty} G_0^-(t-t')|\Psi(t')\rangle \tag{18.5.70}$$

and show that $|\Psi_0'(t)\rangle$ is in fact $|\Psi_{out}(t)\rangle$. Both results are used a little later.

18.6 Integral Equations for Propagators

We have the equations for the propagators

$$\left(i\hbar\frac{\partial}{\partial t} - H_0\right) G_0^\pm(t) = \delta(t)\mathbf{1} \tag{18.6.71}$$

$$\left(i\hbar\frac{\partial}{\partial t} - H\right) G^\pm(t) = \delta(t)\mathbf{1} \tag{18.6.72}$$

and the boundary conditions

$$G_0^+(t) = G^+(t) = 0 \quad \text{for} \quad t < 0 \tag{18.6.73}$$

$$G_0^-(t) = G^-(t) = 0 \quad \text{for} \quad t > 0 \ . \tag{18.6.74}$$

Also, we use the fact that

$$H = H_0 + V \ . \tag{18.6.75}$$

It is possible to use $G_0^\pm$ to derive integral equations for $G^\pm$. In fact simply using (18.6.71) and (18.6.72) we obtain immediately

$$\left(i\hbar\frac{\partial}{\partial t} - H_0\right) G^\pm(t) = \delta(t)\mathbf{1} + VG^\pm(t) \ . \tag{18.6.76}$$

Now, applying the boundary conditions (18.6.73) and (18.6.74) we get

$$G^\pm(t) = G_0^\pm(t) + \int_{-\infty}^{\infty} G_0^\pm(t-t')VG^\pm(t')\,dt' \ . \tag{18.6.77}$$

These equations are integral equations for $G^{\pm}$ in terms of $G_0^{\pm}$. Their Fourier transforms in t play an important role in the time-independent formulation of scattering theory and are known as the *Lippmann-Schwinger* equations.

It is also possible to derive integral equations for the state $|\Psi(t)\rangle$ by using the results for $|\Psi_{in}(t)\rangle$ or $|\Psi_{out}(t)\rangle$ and the Green's functions $G_0^{\pm}(t)$. Thus, we have

$$\left(i\hbar\frac{\partial}{\partial t} - H_0\right)|\Psi(t)\rangle = V|\Psi(t)\rangle . \tag{18.6.78}$$

Hence,

$$|\Psi^{\pm}(t)\rangle = |\Psi_{in,out}(t)\rangle + \int_{-\infty}^{\infty} G_0^{\pm}(t-t')V|\Psi^{\pm}(t')\rangle\, dt' \tag{18.6.79}$$

as is easily checked by differentiation. Also the boundary conditions

$$\lim_{t\to\mp\infty} |\Psi^{\pm}(t)\rangle = |\Psi_{in,out}(t)\rangle \tag{18.6.80}$$

are already included. These equations are sometimes referred to as the *Källén-Yang-Feldman* equations.

The meaning of these equations is as follows. We prepare a state for a scattering experiment. This would normally be a collimated beam moving towards the scattering target. This prepared state is described by $|\Psi_{in}(t)\rangle$ and contains all the information about the incident beam and target at $t = -\infty$. That is, it is labelled with all the appropriate quantum numbers. This state then evolves under the influence of H into the state $|\Psi^{+}(t)\rangle$ which is labelled in exactly the same manner as $|\Psi_{in}(t)\rangle$. In other words, the same set of quantum numbers ($\mathbf{p}$, s^2, s_z, etc.) may be used to label both. Thus, $|\Psi^{+}(t)\rangle$ is a state that arises from a given prepared state $|\Psi_{in}(t)\rangle$ in the remote past. As $|\Psi^{+}(t)\rangle$ continues to evolve, it again becomes a free state $|\Psi_{out}(t)\rangle$ which contains not only the original beam $|\Psi_{in}(t)\rangle$ but also some outgoing scattered wave. One can, in principle, also consider a state specified or selected in the remote future, namely $|\Psi_{out}(t)\rangle$. This state would then arise from the state $|\Psi^{-}(t)\rangle$. In practice this is of course impossible. However, both $|\Psi_{out}(t)\rangle$ and $|\Psi^{-}(t)\rangle$ would also be labelled by the same set of quantum numbers.

18.7 Cross-Sections

In a scattering experiment the quantity actually measured is the number of particles at a given energy scattered by the target into the element of solid angle between Ω and $\Omega + d\Omega$. This number is proportional to what is known as the differential cross-section

$$\frac{d\sigma(\theta,\varphi)}{d\Omega} .$$

Alternatively, suppose the incident beam has a current density $\mathbf{j}_{inc}$ and a fraction of these incident particles are scattered by a potential producing a scattered beam of current density $\mathbf{j}_{scat}$, then dN, the number of scattered particles per unit time that pass through an element of surface area $d\mathbf{s}$, is

$$dN = \mathbf{j}_{scat} \cdot d\mathbf{s} \ . \tag{18.7.81}$$

This number is proportional to $|\mathbf{j}_{inc}|$ and $d\Omega$. The constant of proportionality is defined as the differential cross-section

$$\frac{d\sigma(\theta, \varphi)}{d\Omega} \ .$$

Thus, we have

$$dN = |\mathbf{j}_{inc}| \frac{d\sigma(\theta, \varphi)}{d\Omega} \, d\Omega \tag{18.7.82}$$

where

$$d\Omega = \frac{\hat{n} \cdot d\mathbf{s}}{r^2} \tag{18.7.83}$$

and $\hat{n}$ is a unit vector in the direction of the scattered beam. From this discussion it is clear that the differential cross-section has the dimensions of an area.

One can think of

$$\frac{d\sigma(\theta, \varphi)}{d\Omega}$$

as the effective cross-sectional area of the target particles so that incident particles are scattered by them into the element of solid angle $d\Omega$. Thus, if the incident current consists of J particles per unit area per unit time and if the target, irradiated by the incident beam, contains N_0 particles (scattering centres) then the number of particles dN scattered into the element of solid angle $d\Omega$ per unit time is given by

$$dN = J N_0 \frac{d\sigma(\theta, \varphi)}{d\Omega} \, d\Omega \ . \tag{18.7.84}$$

In practice the observed quantity is dN. The quantities J and N_0 are known from other considerations or measurements and $d\Omega$ is given by the location and effective area of the detector. In this manner one gets a measurement of the differential cross-section $d\sigma/d\Omega$. The total cross-section σ is then defined by

$$\sigma = \int \frac{d\sigma(\theta, \varphi)}{d\Omega} \, d\Omega \tag{18.7.85}$$

where the integral extends over the full solid angle of 4π.

18.8 The Lippmann-Schwinger Equations

In preparation for the time-independent formulation we now start by Fourier transforming the Green's functions $G_0^\pm(t)$ and $G^\pm(t)$. We also call their Fourier transforms $G_0^\pm(E)$ and $G^\pm(E)$ so that the functions are partly defined by their argument. This should not be the source of any confusion. Thus,

$$G_0^\pm(E) = \int_{-\infty}^{\infty} e^{iEt/\hbar}\, G_0^\pm(t)\, dt \tag{18.8.86}$$

and

$$G^\pm(E) = \int_{-\infty}^{\infty} e^{iEt/\hbar}\, G^\pm(t)\, dt \ . \tag{18.8.87}$$

These integrals as they stand are not well defined due to convergence difficulties for large $|t|$. Since $G_0^+(t)$ and $G^+(t)$ both vanish for $t < 0$ we can ensure convergence by giving E a small positive imaginary part $i\epsilon$. Similarly $G_0^-(t)$ and $G^-(t)$ vanish for $t > 0$ and their Fourier transforms are obtained by giving E a small negative imaginary part $-i\epsilon$. Inserting (18.8.86) and (18.8.87) into the differential equations (18.6.71) and (18.6.72) we get

$$G_0^\pm(E) = (E \pm i\epsilon - H_0)^{-1} \tag{18.8.88}$$

$$G^\pm(E) = (E \pm i\epsilon - H)^{-1} \ . \tag{18.8.89}$$

Thus the $\pm i\epsilon$ can be seen to indicate how one is to integrate past the poles in $G_0^\pm(E)$ and $G^\pm(E)$ when transforming back to $G_0^\pm(t)$ and $G^\pm(t)$. They reflect the boundary conditions originally imposed on the time-dependent Green's functions. One can also Fourier transform the state vectors to get

$$|\Psi(E)\rangle = \int_{-\infty}^{\infty} e^{iEt/\hbar} |\Psi(t)\rangle\, dt \ . \tag{18.8.90}$$

as well as

$$|\Psi_{in}(E)\rangle = \int_{-\infty}^{\infty} e^{iEt/\hbar} |\Psi_{in}(t)\rangle\, dt \ . \tag{18.8.91}$$

$$|\Psi_{out}(E)\rangle = \int_{-\infty}^{\infty} e^{iEt/\hbar} |\Psi_{out}(t)\rangle\, dt \ . \tag{18.8.92}$$

If we now Fourier transform equation (18.6.79) and use (18.8.88) we get

$$\begin{aligned} |\Psi^+(E)\rangle &= |\Psi_{in}(E)\rangle + (E + i\epsilon - H_0)^{-1} V |\Psi^+(E)\rangle \\ |\Psi^-(E)\rangle &= |\Psi_{out}(E)\rangle + (E - i\epsilon - H_0)^{-1} V |\Psi^-(E)\rangle \end{aligned} \tag{18.8.93}$$

where $|\Psi^+(E)\rangle$ and $|\Psi^-(E)\rangle$ are states labelled with the same quantum numbers as $|\Psi_{in}(E)\rangle$ and $|\Psi_{out}(E)\rangle$ respectively.

Equations (18.8.93) constitute the Lippmann-Schwinger equations for the wave functions. We can also obtain Lippmann-Schwinger equations for the Green's functions by Fourier transforming equations (18.6.77).

$$\begin{aligned} G^+(E) &= (E + i\epsilon - H_0)^{-1} + (E + i\epsilon - H_0)^{-1} V G^+(E) \\ G^-(E) &= (E - i\epsilon - H_0)^{-1} + (E - i\epsilon - H_0)^{-1} V G^-(E) \ . \end{aligned} \tag{18.8.94}$$

It has become conventional to write

$$(E \pm i\epsilon - H_0)^{-1} \quad \text{as} \quad \frac{1}{E \pm i\epsilon - H_0}$$

and we shall do so freely in the future. From our previous discussion of distributions (Chapter 9) we know that for such functions

$$\frac{1}{E \pm i\epsilon - H_0} = P\frac{1}{E - H_0} \mp i\pi\delta(E - H_0) \ . \tag{18.8.95}$$

This is extended to operators by simply defining

$$\delta(E - H_0)|\Psi^{(0)}(E')\rangle = \delta(E - E')|\Psi^{(0)}(E')\rangle \tag{18.8.96}$$

where

$$H_0|\Psi^{(0)}(E')\rangle = E'|\Psi^{(0)}(E')\rangle \tag{18.8.97}$$

so that $|\Psi^{(0)}(E')\rangle$ is an eigenstate of H_0. Since the $|\Psi^{(0)}(E')\rangle$ form a complete set, this suffices to define $\delta(E - H_0)$. Similarly, $P\frac{1}{E-H_0}$ is defined by

$$P\frac{1}{E - H_0}|\Psi^{(0)}(E')\rangle = P\frac{1}{E - E'}|\Psi^{(0)}(E')\rangle \ . \tag{18.8.98}$$

It is important to notice that these operators are functions of H_0 and are defined by the action on eigenstates of H_0. If they were functions of an operator A they would be defined by the action on eigenstates of A.

18.9 The S-Matrix and the Scattering Amplitude

We have already introduced the S-operator which connects incoming and outgoing states via

$$|\Psi_{in}\rangle = S^\dagger|\Psi_{out}\rangle \tag{18.9.99}$$

$$|\Psi_{out}\rangle = S|\Psi_{in}\rangle \ . \tag{18.9.100}$$

When its matrix elements are evaluated on free states of definite energy, such as plane wave states, it is called the S-matrix and its elements are of considerable interest since they contain all the information for a scattering process. To see this consider a typical scattering experiment. We send a particle down a beam tube. This particle is described by a free state $|\Psi_{in}(t)\rangle = |\Psi_0(\alpha, t)\rangle$ where α is a complete set of labels that are eigenvalues of operators that commute with the *free* Hamiltonian H_0. This state evolves in time into a state $|\Psi^{(+)}(\alpha,t)\rangle$ labelled in exactly the same way. The operators, whose eigenvalues label $|\Psi_0(\alpha,t)\rangle$ and $|\Psi^{(+)}(\alpha,t)\rangle$ do not, however, commute with the full Hamiltonian H. They simply state how $|\Psi^{(+)}(\alpha,t)\rangle$ was prepared in the remote past; they are simply the labels attached to the original particle we sent down the beam tube.

In the remote future this state again evolves into a free state $|\Psi_{out}\rangle$. This state does not carry the same labels as $|\Psi_{in}\rangle$ because it contains an admixture of states due to scattering.

The purpose of any scattering experiment is to measure the probability of finding a given free state $|\Psi_0(\beta,t)\rangle$ in the state that will have evolved in the distant future from the state $|\Psi^{(+)}(\alpha,t)\rangle$. Clearly this is given by

$$\lim_{t\to\infty}\langle\Psi_0(\beta,t)|\Psi^+(\alpha,t)\rangle = \langle\Psi_0(\beta,t)|\Psi_{out}(t)\rangle \tag{18.9.101}$$

and using (18.9.100) this becomes

$$\langle\Psi_0(\beta,t)|S|\Psi_{in}(t)\rangle = \langle\Psi_0(\beta,t)|S|\Psi_0(\alpha,t)\rangle\ . \tag{18.9.102}$$

Thus, as stated, the elements of the S-matrix contain all the information obtained in a scattering experiment.

In order to obtain a useful formula for these matrix elements we view the scattering experiment in a somewhat different fashion.

The state $|\Psi_{out}(t)\rangle$ must arise from the state $|\Psi^-(\beta,t)\rangle$. Again the labels β are the eigenvalues of a complete set of observables that commute with H_0. It is, of course impossible to prepare such a state $|\Psi_{out}(\beta,t)\rangle$. Nevertheless we can now think of a scattering experiment as a means to measure the probability of finding in the state $|\Psi^-(\beta,t)\rangle$ the particle described by $|\Psi^+(\alpha,t)\rangle$ and prepared in the remote past as $|\Psi_0(\alpha,t)\rangle$. The corresponding probability amplitude must coincide with the one obtained above. Hence, we find

$$\langle\Psi_0(\beta,t)|S|\Psi_0(\alpha,t)\rangle = \langle\Psi^-(\beta,t)|\Psi^+(\alpha,t)\rangle\ . \tag{18.9.103}$$

This result will be used immediately in conjunction with the Lippmann-Schwinger equations to obtain a neat formula for the S-matrix elements. For this purpose we consider eigenstates of the Hamiltonian H_0. Thus our states are labelled $|\Psi_0(E)\rangle$ where temporarily we suppress the additional labels α.

As a first step we write a formal solution of the Lippmann-Schwinger equations

$$\begin{aligned}(E-H)|\Psi^\pm\rangle &= (E-H_0)|\Psi_0\rangle\\ &= (E-H+V)|\Psi_0\rangle\\ &= (E-H)|\Psi_0\rangle + V|\Psi_0\rangle\ .\end{aligned} \tag{18.9.104}$$

Thus,

$$|\Psi^\pm\rangle = |\Psi^0\rangle + (E-H\pm i\epsilon)^{-1}V|\Psi_0\rangle \tag{18.9.105}$$

where $|\Psi_0\rangle$ is a free state of energy E corresponding to an incoming or outgoing wave depending on the appropriate initial condition specified by $\pm i\epsilon$. The states $|\Psi^{\pm}\rangle$ are labelled by the same set of quantum numbers as $|\Psi_0\rangle$. The labels $\pm$ er to the $\pm i\epsilon$ and simply mean that for $|\Psi^+\rangle$, the state $|\Psi_0\rangle$ is an incoming wave whereas for $|\Psi^-\rangle$ it is an outgoing wave. The function $|\Psi_0\rangle$ is the same function in either case; it is just a solution of the free Schrödinger equation. Note that equation (18.9.105) is purely a *formal* solution since we must still evaluate the operator $(E - H \pm i\epsilon)^{-1}$ and from our definition of such expressions this require a knowledge of the eigenfunctions of H. Thus, we have not really succeeded in solving the Lippmann-Schwinger equations, we have simply rewritten them.

As we saw (equation (18.9.103)) the S-matrix elements S_{fi} are specified by

$$\begin{aligned} S_{fi} &= \langle\Psi^-(E_f)|\Psi^+(E_i)\rangle \\ &= \langle\Psi^+(E_f)|\Psi^+(E_i)\rangle + \left(\langle\Psi^-(E_f) - \langle\Psi^+(E_f)\right)|\Psi^+(E_i)\rangle \end{aligned} \tag{18.9.106}$$

where we have chosen energy eigenstates for $|\Psi_0\rangle$ and are still suppressing all other labels. Now using our formal solutions (18.9.105) we can write

$$\langle\Psi^-(E_f)| - \langle\Psi^+(E_f)| = \langle\Psi^0(E_f)|\left(\frac{1}{E_f - H + i\epsilon}V - \frac{1}{E_f - H - i\epsilon}V\right). \tag{18.9.107}$$

Also the kets $|\Psi^{\pm}(E_f)\rangle$ are assumed orthonormal so that

$$\langle\Psi^{\pm}(E_f)|\Psi^{\pm}(E_i)\rangle = \delta_{fi} \tag{18.9.108}$$

where

$$\delta_{fi} = \delta(E_f - E_i)\delta_{\alpha,\beta}\ . \tag{18.9.109}$$

Here α, β specify all the other quantum numbers, besides energy, which we have so far suppressed and which are required to completely label $|\Psi^+(E)\rangle$. We therefore obtain the relation

$$\begin{aligned} S_{fi} &= \delta_{fi} + \langle\Psi^0(E_f)|\left(V\frac{1}{E_f - H + i\epsilon} - V\frac{1}{E_f - H - i\epsilon}\right)|\Psi^+(E_i)\rangle \\ &= \delta_{fi} + \left(P\frac{1}{E_f - E_i} - i\pi\delta(E_f - E_i)\right. \\ &- \left.P\frac{1}{E_f - E_i} - i\pi\delta(E_f - E_i)\right)\langle\Psi^0(E_f)|V|\Psi^+(E_i)\rangle \end{aligned} \tag{18.9.110}$$

where we have used the relations (18.8.95), (18.8.96), and (18.8.98). Thus, finally

$$S_{fi} = \delta_{fi} - 2\pi i\delta(E_f - E_i)\langle\Psi^0(E_f)|V|\Psi^+(E_i)\rangle \tag{18.9.111}$$

or somewhat more concisely

$$S_{fi} = \delta_{fi} - 2\pi i\delta(E_f - E_i)T_{fi} \tag{18.9.112}$$

where

$$T_{fi} = \langle\Psi^0(E_f)|V|\Psi^+(E_i)\rangle \quad \text{for} \quad E_f = E_i \tag{18.9.113}$$

is the *T-matrix element on the energy shell.* Multiplied by $-1/4\pi$ it is called the *scattering amplitude* $f(k,\theta)$. Aside from the conservation of energy and other quantum numbers specified by the two delta functions, T_{fi} contains all the information about a scattering process.

If originally we had replaced $|\Psi^+(E_i)\rangle$ by

$$|\Psi^-(E_i)\rangle + |\Psi^+(E_i)\rangle - |\Psi^-(E_i)\rangle$$

then by exactly similar steps we would have arrived at the relationship

$$S_{fi} = \delta_{fi} - 2\pi i\delta(E_f - E_i)\langle\Psi^-(E_f)|V|\Psi^0(E_i)\rangle \tag{18.9.114}$$

so that we also have

$$T_{fi} = \langle\Psi^-(E_f)|V|\Psi^0(E_i)\rangle \quad \text{for} \quad E_f = E_i \ . \tag{18.9.115}$$

In the next chapter we develop a systematic formulation of scattering theory starting from the time-independent Schrödinger equation. During this process we rederive some of the results obtained here. In this way we hope, not only to emphasize the more important results, but also to elucidate the concepts involved. In practice, computations commence almost always with the time-independent formalism.

18.10 Problems

18.1 A proton beam producing a current of 5×10^{-9} amps is incident on a target of copper. Assume the target thickness is such that the areal density is 0.2 mg/cm^2. The detector has an area of 0.5 cm^2, normal to the scattered beam, and is 20 cm from the target. If 10 protons are counted by the detector every second at a particular angle, calculate the differential cross-section for protons scattering off copper at that angle.

18.2 Use the expressions (18.5.47) and (18.5.48) and evaluate the matrix elements $\langle \mathbf{p}|G_0^{\pm}(t)|\mathbf{k}\rangle$ where $|\mathbf{p}\rangle$, $|\mathbf{k}\rangle$ are free particle states of momentum $\mathbf{p}$ and $\mathbf{k}$ respectively.

18.3 Calculate the Fourier transform of the distribution $\theta(t)$.
Hint:

$$\theta(\omega) = \lim_{\epsilon\to 0+}\int_0^\infty e^{i(\omega - i\epsilon)t}\,dt \ .$$

The limit here is to be understood in the sense of distributions.

18.4 Assume that V is independent of time and use (18.6.77) to obtain an equation for the Fourier transform $G^{\pm}(\omega)$ of $G^{\pm}(t)$ in terms of the Fourier transforms of $G_0^{\pm}(t)$. Write a formal solution for $G^{\pm}(\omega)$.

18.5 In equation (18.9.113) approximate $|\Psi^+(E_i)\rangle$ by a free particle state. If V is a screened Coulomb potential

$$V = -V_0 \frac{e^{-\mu r}}{r}$$

calculate the scattering amplitude. The approximation used is known as the first Born approximation.

Bibliography

[18.1] There are several books devoted exclusively to scattering theory. Two very detailed books are:
M.L. Goldberger and K.M. Watson, *Collision Theory* - John Wiley and Sons, Inc., New York (1964),
N.F. Mott and H.S.W. Massey, *The Theory of Atomic Collisions* - Oxford University Press, London 2nd edition (1948).

[18.2] A book treating also classical scattering theory as well as quantum scattering theory is R.G. Newton, *Scattering Theory of Waves and Particles* - McGraw-Hill Book Co., New York (1966).
This book contains many interesting worked examples as well as a wealth of useful information.

Chapter 19

Scattering Theory - Time Independent

19.1 Introduction

In the previous chapter we gave a time-dependent formulation of scattering. By Fourier transforming these quations we obtained a time-independent set of equations known as the Lippmann-Schwinger equations. In this chapter we start from the time-independent Schrödinger equation and rederive these results along the way. This chapter is independent of all of Chapter 18, except section 18.7, and may be studied before Chapter 18. The results obtained here provide an efficient means for computing differential cross-sections. To recall the definition of cross-section it is worthwhile to read or re-read section 18.7.

19.2 The Scattering Amplitude

We start with the time-independent Schrödinger equation for a one-body problem where

$$H = \frac{\mathbf{p}^2}{2m} + V(r) \tag{19.2.1}$$

and m may be the reduced mass of a two-particle system in which case $V(r)$ is the potential between the two particles. It is usual in potential scattering to introduce

$$k^2 = \frac{2mE}{\hbar^2} \qquad U(r) = \frac{2mV(r)}{\hbar^2} \ . \tag{19.2.2}$$

In that case the time-independent Schrödinger equation corresponding to the Hamiltonian H becomes:

$$\left(\nabla^2 + k^2\right) \psi(\mathbf{r}) = U(r)\psi(\mathbf{r}) \ . \tag{19.2.3}$$

In order to have a mathematically simple scattering problem it is desirable that $U(r) \to 0$ sufficiently rapidly as $r \to \infty$. We assume this to be the case. [1] With these conditions satisfied we can look for solutions of (19.2.3) that consist of a given incoming wave say $e^{i\mathbf{k}\cdot\mathbf{r}}$ plus a scattered wave $\psi_{scat}(\mathbf{r})$ so that

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} + \psi_{scat}(\mathbf{r}) \ . \tag{19.2.4}$$

The current due to the incident beam is

$$\mathbf{j}_{inc} = \frac{\hbar\mathbf{k}}{m} \tag{19.2.5}$$

and corresponds to a flux of particles moving in the $\mathbf{k}$ direction with uniform momentum $\hbar\mathbf{k}$. If we consider a very large sphere of radius R with centre at the scattering centre (potential) namely $r = 0$, then the flux through the surface of that sphere due to the scattered wave is

$$R^2 \oint \left(\mathbf{j}_{scat} \cdot \frac{\mathbf{r}}{r}\right)_R d\Omega$$

where $(\mathbf{j}_{scat} \cdot \mathbf{r}/r)_R$ is the radial component of the current due to the scattered beam, evaluated at $r = R$. This is given by

$$(\mathbf{j}_{scat} \cdot \mathbf{r}/r)_R = \frac{\hbar}{2im}\left(\psi^*_{scat}\frac{\partial\psi_{scat}}{\partial r} - \psi_{scat}\frac{\partial\psi^*_{scat}}{\partial r}\right)\bigg|_{r=R} \ . \tag{19.2.6}$$

If the scattered flux through the surface of the sphere is to tend to a constant value independent of R as $R \to \infty$, we must have

$$\Im\left(\psi^*_{scat}\frac{\partial\psi_{scat}}{\partial r}\right)\bigg|_{r=R} \to \frac{\text{function of } k\&\theta}{R^2} \quad \text{for} \quad R \to \infty \ . \tag{19.2.7}$$

This means that

$$\lim_{R\to\infty} \psi_{scat}(R) = -\frac{1}{4\pi}T(\mathbf{k},\mathbf{k}')\frac{e^{ikR}}{R} = f(k,\theta)\frac{e^{ikR}}{R} \tag{19.2.8}$$

where $\mathbf{k}$ is the incident momentum, and $\mathbf{k}'$ is the scattered or outgoing momentum and θ is the angle between $\mathbf{k}$ and $\mathbf{k}'$. The quantity $f(k,\theta)$ is called the scattering amplitude. The flux of scattered particles through an element of solid angle $d\Omega$ is therefore given by

$$\begin{aligned} R^2 d\Omega\, \frac{\hbar k}{m}|f(k,\theta)|^2 \,\frac{1}{R^2} &= \frac{\hbar k}{m}|f(k,\theta)|^2 d\Omega \\ &= |\mathbf{j}_{inc}||f(k,\theta)|^2 d\Omega \ . \end{aligned} \tag{19.2.9}$$

Hence, using the definition of $d\sigma/d\Omega$, equation (18.7.82), we see that the differential cross-section is given by

$$\frac{d\sigma}{d\Omega} = |f(k,\theta)|^2 \ . \tag{19.2.10}$$

Since the experimentally accessible quantity we are interested in is $d\sigma/d\Omega$ our task is reduced to computing the scattering amplitude $f(k,\theta)$.

---

[1] $r^2V(r) \to 0$ for $r \to \infty$ is sufficient.

19.3 Green's Functions

In order to find a solution of the form (19.2.4) we convert the Schrödinger equation (19.2.3) to an integral equation. To do this we must determine the Green's function satisfying the relation

$$\left(\nabla^2 + k^2\right) G_0(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') \ . \tag{19.3.11}$$

This is analogous to solving for an operator $G_0(E)$ using

$$(-H_0 + E)G_0(E) = 1 \ . \tag{19.3.12}$$

Now for E complex, say z, the operator-valued function

$$G_0(z) = (z - H_0)^{-1} \tag{19.3.13}$$

is well defined [2] since the eigenfunctions $\psi^{(0)}(E)$ of H_0 form a complete set

$$H_0\psi^{(0)}(E) = E\psi^{(0)}(E) \ . \tag{19.3.14}$$

Thus,

$$G_0\psi^{(0)}(E) = \frac{1}{z-E}\,\psi^{(0)}(E) \ . \tag{19.3.15}$$

When z approaches real values this operator develops poles in z at the eigenvalues and must be defined in the sense of distributions. We examine this point more explicitly starting from (19.3.11). If we Fourier transform (19.3.11) using

$$G_0(\mathbf{r},\mathbf{r}') = \frac{1}{(2\pi)^3}\int e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}\tilde{G}_0(\mathbf{q})\, d^3q \tag{19.3.16}$$

and use the integral representation of the δ function

$$\delta(\mathbf{r}-\mathbf{r}') = \frac{1}{(2\pi)^3}\int e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}\, d^3q \tag{19.3.17}$$

then,

$$\tilde{G}_0(\mathbf{q}) = \frac{1}{k^2-q^2} \ . \tag{19.3.18}$$

Consequently

$$G_0(\mathbf{r},\mathbf{r}') = \frac{1}{(2\pi)^3}\int \frac{e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}}{k^2-q^2}\, d^3q \ . \tag{19.3.19}$$

Since the integrand has poles on the path of integration, we must specify how this integral is to be defined. Thus, it can be defined as a principal value integral by omitting the poles or differently by including a portion of the residue due to

---

[2] It is called the resolvent of H_0. See also problem (6.6).

each pole. Each of these integrations gives rise to a different Green's functions. We are particularly interested in two of these Green's functions, namely

$$G_0^{\pm}(\mathbf{r},\mathbf{r}') = \lim_{\epsilon\to 0+} \frac{-1}{(2\pi)^3} \int \frac{e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}}{q^2 - (k^2 \pm i\epsilon)}\, d^3q \ . \tag{19.3.20}$$

We can immediately integrate over the azimuthal angle if we choose the q_z-axis aligned with $\mathbf{r}-\mathbf{r}'$ so that

$$\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}') = q|\mathbf{r}-\mathbf{r}'|\cos\theta \ . \tag{19.3.21}$$

This choice for the z-direction of the integration variable is a trick that can be frequently used to evaluate three-dimensional Fourier integrals. The integral in (19.3.20) now reduces to

$$G_0^{\pm}(\mathbf{r},\mathbf{r}') = \lim_{\epsilon\to 0+} \frac{-2\pi}{(2\pi)^3} \int_0^{\infty} \frac{q^2\,dq}{q^2 - k^2 \mp i\epsilon} \int_{-1}^{1} e^{iq|\mathbf{r}-\mathbf{r}'|u}\, du \tag{19.3.22}$$

where we have integrated over φ and made the substitution $u = \cos\theta$. The integration over u is now easily performed to yield

$$\begin{aligned} G_0^{\pm}(\mathbf{r},\mathbf{r}') &= \lim_{\epsilon\to 0+} \frac{-1}{(2\pi)^2}\frac{1}{i|\mathbf{r}-\mathbf{r}'|} \int_0^{\infty} \frac{q^2\,dq}{q^2 - k^2 \mp i\epsilon} \frac{e^{iq|\mathbf{r}-\mathbf{r}'|} - e^{-iq|\mathbf{r}-\mathbf{r}'|}}{q} \\ &= \lim_{\epsilon\to 0+} \frac{i}{(2\pi)^2}\frac{1}{|\mathbf{r}-\mathbf{r}'|} \int_{-\infty}^{\infty} \frac{e^{iq|\mathbf{r}-\mathbf{r}'|}}{q^2 - k^2 \mp i\epsilon}\, q\,dq \ . \end{aligned} \tag{19.3.23}$$

To evaluate these integrals we now close the contour with a half circle in the upper half of the q-plane so that the exponential is damped. The poles in the integrant are located either at $\pm(k+i\epsilon)$ or at $\pm(k-i\epsilon)$. In either case we obtain only the residue from the pole in the upper half plane. Thus, we get

$$G_0^{\pm}(\mathbf{r},\mathbf{r}') = \frac{-1}{4\pi}\frac{e^{\pm ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \ . \tag{19.3.24}$$

If we remember that time dependence gives us a factor $e^{i\omega t}$ then we see that G^+ corresponds to a wave travelling away from the potential centered at $r=0$ (an outgoing wave) and G^- corresponds to an incoming wave.

The expressions in (19.3.24) are simply the configuration space expressions for the operators

$$G_0^{\pm}(E) = \frac{1}{E \pm i\epsilon - H_0} \tag{19.3.25}$$

that we considered before. It is now a simple matter to rewrite the time-independent Schrödinger equation as an integral equation. The boundary conditions are incorporated in the $\pm i\epsilon$. In fact the equation is given by

$$\Psi^{\pm}(\mathbf{r}) = \Psi^0(\mathbf{r}) + \int G_0^{\pm}(\mathbf{r},\mathbf{r}')U(\mathbf{r}')\Psi^{\pm}(\mathbf{r}')\, d^3r' \tag{19.3.26}$$

where Ψ^0 is a solution of the free Schrödinger equation. The fact that $\Psi^\pm(\mathbf{r})$ satisfy the correct boundary conditions is clear from the short range nature of $U(\mathbf{r})$ which allows us to take the limit $r \to \infty$ under the integral sign. To see that $\Psi^\pm(\mathbf{r})$ satisfy equation (19.2.3) we need only apply the operator $\nabla^2 + k^2$ to both sides and use equation (19.3.26). We then find

$$\begin{aligned}(\nabla^2 + k^2)\Psi^\pm(\mathbf{r}) &= 0 + \int \delta(\mathbf{r} - \mathbf{r}')U(\mathbf{r}')\Psi^\pm(\mathbf{r}')\, d^3r' \\ &= U(\mathbf{r})\Psi^\pm(\mathbf{r}) \qquad (19.3.27)\end{aligned}$$

as required.

Thus, the solutions $\Psi^\pm$ corresponding to $G_0^\pm$ do indeed satisfy equation (19.3.26). Going back to the Schrödinger equation in the form

$$(E - H_0)|\Psi^0(E)\rangle = V|\Psi^0(E)\rangle \qquad (19.3.28)$$

we get

$$|\Psi^\pm(E)\rangle = |\Psi^0(E)\rangle + \frac{1}{E - H_0 \pm i\epsilon} V|\Psi^\pm(E)\rangle \; . \qquad (19.3.29)$$

Thus, (19.3.26) is simply (19.3.29) written out in configuration space. In either case the equations are known as the Lippmann-Schwinger equations.

To show the equivalence of (19.3.26) and (19.3.29) we simply transcribe (19.3.29) into the configuration space representation using the relations

$$\Psi^\pm(\mathbf{r}) = \langle \mathbf{r}|\Psi^\pm(E)\rangle \qquad (19.3.30)$$

and

$$\begin{aligned}\Psi^0(\mathbf{r}) &= \langle \mathbf{r}|\Psi^0(E)\rangle \\ &= \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{p}\cdot\mathbf{r}} \\ &= \langle \mathbf{r}|\mathbf{p}\rangle \qquad (19.3.31)\end{aligned}$$

where

$$E = \frac{\hbar^2 \mathbf{p}^2}{2m} \; . \qquad (19.3.32)$$

Then, (19.3.29) becomes

$$\begin{aligned}\Psi^\pm(\mathbf{r}) &= \Psi^0(\mathbf{r}) + \langle \mathbf{r}|\frac{1}{E - H_0 \pm i\epsilon} V|\Psi^\pm(E)\rangle \\ &= \Psi^0(\mathbf{r}) + \int d^3k\, d^3q\, d^3r'\, d^3r'' \langle \mathbf{r}|\mathbf{k}\rangle \cdot \\ &\quad \langle \mathbf{k}|\frac{1}{E - H_0 \pm i\epsilon}|\mathbf{q}\rangle\langle \mathbf{q}|\mathbf{r}'\rangle\langle \mathbf{r}'|V|\mathbf{r}''\rangle\langle \mathbf{r}''|\Psi^\pm(E)\rangle \qquad (19.3.33)\end{aligned}$$

where we now use the fact that V is local so that

$$\langle \mathbf{r}'|V|\mathbf{r}''\rangle = V(\mathbf{r}')\delta(\mathbf{r}' - \mathbf{r}'') \; . \qquad (19.3.34)$$

Next, writing $E = \hbar^2 \mathbf{p}^2/2m$, the integral over k becomes

$$\frac{2m}{\hbar^2}\frac{1}{(2\pi)^3}\int \frac{d^3k}{p^2-k^2\pm i\epsilon}\, e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} = \frac{2m}{\hbar^2} G_0^{\pm}(\mathbf{r}-\mathbf{r}') \ . \tag{19.3.35}$$

Putting this back into (19.3.33) and recalling that

$$U(\mathbf{r}) = \frac{2m}{\hbar^2} V(\mathbf{r}) \tag{19.3.36}$$

we immediately obtain (19.3.26).

19.4 The Born Approximation

The Lippmann-Schwinger equations provide an immediate approximation technique when the interaction V is small compared to the energy E. Thus, it becomes more exact as the energy increases. To get the Born [19.3] series [3] one simply iterates equation (19.3.29) to get

$$|\Psi^{\pm}(E)\rangle = \sum_{n=0}^{\infty}\left(\frac{1}{E-H_0\pm i\epsilon}V\right)^n |\Psi^0(E)\rangle \ . \tag{19.4.37}$$

The usefulness of this series is that for small V or large E it may be truncated after only a few terms. In fact, in practice one frequently keeps only the first non-trivial term. This is known as the *first Born* or simply *Born* approximation. Written out it reads

$$|\Psi^{\pm}(E)\rangle = |\Psi^0(E)\rangle + \frac{1}{E-H_0\pm i\epsilon}V|\Psi^0(E)\rangle \ . \tag{19.4.38}$$

Going over to configuration space this becomes for Ψ^+, the wave that contains the scattered or outgoing wave,

$$\Psi^+(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{2m}{\hbar^2}\int G_0^+(\mathbf{r}-\mathbf{r}')V(\mathbf{r}')\,\Psi^0(\mathbf{r}')\,d^3r' \ . \tag{19.4.39}$$

We could have obtained this directly, of course, by just iterating equation (19.3.26) once and dropping all other terms.

Substituting for $\Psi^0(\mathbf{r}')$ as well as $G_0^+(\mathbf{r}-\mathbf{r}')$ we obtain

$$\Psi^+(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} - \frac{2m}{\hbar^2}\frac{1}{4\pi}\int \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}V(\mathbf{r}')\,e^{i\mathbf{k}\cdot\mathbf{r}'}\,d^3r' \ . \tag{19.4.40}$$

To get the scattering amplitude we require the asymptotic solution for large r. Now,

$$\begin{aligned} |\mathbf{r}-\mathbf{r}'|^2 &= r^2 - 2rr'\cos\alpha + r'^2 \\ &\to r^2\left(1-\frac{2r'}{r}\cos\alpha\right) \quad \text{for} \quad r\to\infty \end{aligned} \tag{19.4.41}$$

---

[3] Mathematicians call this the Neumann series. See Chapter 2 of reference [19.1].

where

$$\cos\alpha = \frac{\mathbf{r}\cdot\mathbf{r}'}{rr'} \quad . \tag{19.4.42}$$

Therefore,

$$|\mathbf{r}-\mathbf{r}'| \to r - r'\cos\alpha \quad \text{for} \quad r\to\infty \, . \tag{19.4.43}$$

Similarly,

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} \to \frac{1}{r} + \frac{r'\cos\alpha}{r^2} \quad \text{for} \quad r\to\infty \, . \tag{19.4.44}$$

Since we are only interested in the solution which behaves asymptotically as e^{ikr}/r (see the discussion of section 2) we can drop the higher order term in the denominator at this stage. We must, however retain it in the argument of the exponential since the exponential varies more rapidly. Thus, we have

$$\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \to \frac{e^{ik(r-r'\cos\alpha)}}{r} \quad \text{for} \quad r\to\infty \, . \tag{19.4.45}$$

Since $V(r)$ is assumed short range, the integral in equation (19.4.40) converges uniformly and we are justified in using the asymptotic form in the integrand. Thus,

$$\Psi_{\text{Born}} \to e^{i\mathbf{k}\cdot\mathbf{r}} - \frac{2m}{\hbar^2}\frac{1}{4\pi}\frac{e^{ikr}}{r}\int e^{i\mathbf{k}\cdot\mathbf{r}'-ikr'\cos\alpha}\, V(r')\, d^3r' \quad \text{for} \quad r\to\infty \, . \tag{19.4.46}$$

If we now call the scattered momentum $\mathbf{k}'$ such that

$$\mathbf{k}'\cdot\mathbf{r}' = kr'\cos\alpha \tag{19.4.47}$$

then the argument of the exponential becomes $i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}'$. We further define the *change in momentum* or *momentum transfer* $\mathbf{q}$ as

$$\mathbf{q} = \mathbf{k}-\mathbf{k}' \quad . \tag{19.4.48}$$

Then calling θ the scattering angle (angle between $\mathbf{k}$ and $\mathbf{k}'$) we have, since the scattering is elastic (conserves energy), that

$$|\mathbf{k}| = |\mathbf{k}'| = k \quad . \tag{19.4.49}$$

$$q^2 = k^2 - 2k^2\cos\theta + k^2 = 2k^2(1-\cos\theta) = 4k^2\sin^2(\theta/2) \quad . \tag{19.4.50}$$

This gives

$$q = 2k|\sin(\theta/2)| \quad . \tag{19.4.51}$$

The Born term now yields

$$\Psi_{\text{Born}} \to e^{i\mathbf{k}\cdot\mathbf{r}} - \frac{1}{4\pi}\frac{e^{ikr}}{r}\int e^{i\mathbf{q}\cdot\mathbf{r}'}\, U(r')\, d^3r' \quad \text{for} \quad r\to\infty \, . \tag{19.4.52}$$

Or defining the Fourier transform $\tilde{U}(\mathbf{q})$ of $U(\mathbf{r}')$ by

$$\tilde{U}(\mathbf{q}) = \int e^{i\mathbf{q}\cdot\mathbf{r}'} U(\mathbf{r}') d^3r' \tag{19.4.53}$$

we obtain

$$\Psi_{\text{Born}} \to e^{i\mathbf{k}\cdot\mathbf{r}} - \frac{1}{4\pi}\frac{e^{ikr}}{r}\tilde{U}(\mathbf{q}) \quad \text{for} \quad r \to \infty\,. \tag{19.4.54}$$

Comparing this with (19.2.8) we see that in the first order Born approximation the scattering amplitude $f(k,\theta)$ is given by

$$f(k,\theta) = -\frac{1}{4\pi}\tilde{U}(\mathbf{k}-\mathbf{k}')\,. \tag{19.4.55}$$

19.5 The Yukawa Potential

As an example of the application of the Born approximation we consider the Yukawa or screened Coulomb potential

$$V(r) = -V_0\,\frac{e^{-\mu r}}{r}\;. \tag{19.5.56}$$

In this case

$$U(r) = -\frac{2mV_0}{\hbar^2}\int e^{i\mathbf{q}\cdot\mathbf{r}}\,\frac{e^{-\mu r}}{r}\,d^3r\,. \tag{19.5.57}$$

Performing the angular integrations, as before, we now get

$$\tilde{U}(\mathbf{q}) = -\frac{2mV_0}{\hbar^2}\frac{2\pi}{iq}\int_0^\infty \left[e^{-r(\mu-iq)} - e^{-r(\mu+iq)}\right]dr\,. \tag{19.5.58}$$

Hence,

$$\tilde{U}(\mathbf{q}) = -\frac{2mV_0}{\hbar^2}\frac{4\pi}{q^2+\mu^2}\,. \tag{19.5.59}$$

Thus,

$$f(k,\theta) = \frac{2mV_0}{\hbar^2}\frac{1}{q^2+\mu^2}\,. \tag{19.5.60}$$

Substituting this into (19.2.10) and using (19.4.51) we obtain the differential cross-section

$$\frac{d\sigma}{d\Omega} = \frac{1}{2\pi}\frac{d\sigma}{d\cos\theta} = \frac{4m^2V_0^2}{\hbar^4}\frac{1}{(4k^2\sin^2(\theta/2)+\mu^2)^2}\,. \tag{19.5.61}$$

If $V_0 = Ze^2$ and $\mu = 0$ this becomes

$$\frac{1}{2\pi}\frac{d\sigma}{d\cos\theta} = \frac{4m^2Z^2e^4}{16\hbar^4k^4\sin^4(\theta/2)}\,. \tag{19.5.62}$$

Recalling that $(\hbar^2k^2)/2m = E$ we get

$$\frac{1}{2\pi}\frac{d\sigma}{d\cos\theta} = \frac{Z^2e^4}{16E^2\sin^4(\theta/2)} \tag{19.5.63}$$

which is just the classical Rutherford cross-section for the scattering of an electron of charge e by a Coulomb potential of charge Ze. Although the formula (19.5.63) is correct its derivation is not entirely correct since for $\mu = 0$ the potential (19.5.56) is no longer short-ranged. This trick of replacing the Coulomb potential $1/r$ by $e^{-\mu r}/r$ and letting $\mu \to 0$ at the end permits us to use our standard scattering theory without worrying about the long-range nature of the Coulomb potential. The pure Coulomb potential leads to logarithmically oscillating phase contributions ([18.2] (section 14.6) in the wave function ψ^+. In fact,

$$\psi^+ \sim e^{i[kz-n\ln k(r-z)]} + f(k,\theta)\frac{e^{i[kr+n\ln 2kr]}}{r} + \text{higher order terms}$$

where $n = (Ze^2m)/\hbar^2$. In this case one treats $f(k,\theta)$ as the scattering amplitude. These complications can be avoided by using the "adiabatic switching off" of the Coulomb potential for long range with the factor $e^{-\mu r}$ and then letting $\mu \to 0$.

It is rather fortunate that for the case of Coulomb scattering the quantum mechanical and classical cross-sections agree, for otherwise Rutherford's famous experiments would not have yielded the simple interpretation of a tiny nuclear core inside an atom that they did yield. The results of these experiments were of paramount importance in the development of quantum mechanics since they were the basis of Bohr's model as well as Schrödinger's computation of the spectrum of the hydrogen atom.

The fact that the scattering amplitude and hence the differential cross-section are independent of the azimuthal angle φ is not simply a coincidence but is a consequence of the spherical symmetry of the potential. Thus, just as for classical scattering, a spherically symmetric potential gives rise to a differential cross-section independent of the azimuthal angle. We shall begin to exploit this in the next sections. As a first step we develop some more mathematical machinery.

19.6 Free Particle in Spherical Coordinates

In section 10.2 we had occasion to consider the Schrödinger equation in spherical coordinates. The solutions involved spherical Bessel functions. We now re-examine this problem systematically and develop some properties of these Bessel functions. We start with the free Schrödinger equation

$$(\nabla^2 + k^2)\psi = 0 \ . \tag{19.6.64}$$

The solution can be expanded in spherical harmonics

$$\psi(\mathbf{r}) = \sum_{l,m} R_l(r)Y^*_{l,m}(\hat{k})Y_{l,m}(\hat{r}) \tag{19.6.65}$$

where $\hat{k}$, $\hat{r}$ represent the θ, φ directions of $\mathbf{k}$ and $\mathbf{r}$ respectively. Substituting (19.6.65) into (19.6.64) we are led to the radial equation

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} + k^2\right] R_l(r) = 0 \ . \tag{19.6.66}$$

As stated in section 10.2, the solutions of this equation are the spherical Bessel functions $j_l(kr)$ and $n_l(kr)$. In addition to these it is convenient to introduce two additional solutions corresponding to Hankel functions of the first and second kind namely

$$h_l^{(1)}(kr) = j_l(kr) + in_l(kr) \tag{19.6.67}$$

$$h_l^{(2)}(kr) = j_l(kr) - in_l(kr) \ . \tag{19.6.68}$$

To explore the properties of these functions we derive integral representations for them. We also replace kr by x and $R_l(r)$ by $Z_l(x)$ in order to simplify the notation. Thus, (19.6.66) becomes

$$\left[\frac{d^2}{dx^2} + \frac{2}{x}\frac{d}{dx} + 1 - \frac{l(l+1)}{x^2}\right] Z_l(x) = 0 \ . \tag{19.6.69}$$

To derive the integral representations we look for a solution of the form

$$Z_l(x) = x^\lambda \int_a^b e^{xu} f(u)\, du \tag{19.6.70}$$

where we specify the limits a, b later. Substituting this into (19.6.69) we obtain

$$x^{\lambda-2} \int_a^b e^{xu} f(u) \left[\lambda(\lambda+1) - l(l+1) + 2(\lambda+1)xu + (u^2+1)x^2\right] du = 0 . \tag{19.6.71}$$

We now choose $\lambda(\lambda+1) = l(l+1)$ so that

$$\lambda = l \quad \text{or} \quad \lambda = -(l+1) \ . \tag{19.6.72}$$

We then obtain from (19.6.71), after cancelling a factor x

$$x^{\lambda-2} \int_a^b f(u) \left[2(\lambda+1)u + (u^2+1)\frac{d}{du}\right] e^{xu}\, du = 0 \ . \tag{19.6.73}$$

Next we rewrite the second term as a total differential by using the identity

$$\frac{d}{du}\left[f(u)(u^2+1)\, e^{xu}\right] = e^{xu}\, \frac{d}{du}\left[f(u)(u^2+1)\right] + f(u)(u^2+1)\frac{d}{du}e^{xu} . \tag{19.6.74}$$

This yields

$$\begin{aligned} &\int_a^b e^{xu} \left\{2(\lambda+1)uf(u) - \frac{d}{du}\left[f(u)(u^2+1)\right]\right\} du \\ &+ \int_a^b \frac{d}{du}\left[f(u)(u^2+1)\, e^{xu}\right] du = 0 \ . \end{aligned} \tag{19.6.75}$$

We now make both integrals vanish separately: the first by causing the integrand to vanish and the second by an appropriate choice of the limits of integration. The first condition gives

$$\frac{d}{du}\left[f(u)(u^2+1)\right] = 2(\lambda+1)uf(u) \; . \tag{19.6.76}$$

The solution of this equation is

$$f(u) = f(0)(u^2+1)^{\lambda} \; . \tag{19.6.77}$$

To get the second condition we integrate the last term in (19.6.75) to get

$$f(b)(b^2+1)\,e^{xb} - f(a)(a^2+1)\,e^{xa} = 0 \; . \tag{19.6.78}$$

Since x is arbitrary this requires both $f(a)(a^2+1)\,e^{xa}$ and $f(b)(b^2+1)\,e^{xb}$ to vanish. After substituting for f from (19.6.77) this reduces to

$$e^{xa}(a^2+1)^{\lambda+1} = 0 \tag{19.6.79}$$

and

$$e^{xb}(b^2+1)^{\lambda+1} = 0 \; . \tag{19.6.80}$$

For $\lambda = l \geq 0$ this is easily accomplished by choosing $a = -i$, $b = i$. Another possibility for $0 \leq \arg x \leq \pi/2$ is to take one of the limits, that is either a or b, as $\infty e^{i\alpha}$ where $\pi/2 \leq \alpha \leq \pi$. Thus, there are three possibilities:

$$(-i,\, i) \qquad (\infty e^{i\alpha},\, i) \qquad (-i,\, \infty e^{i\alpha}) \; .$$

Each of these limits gives rise to a different function $Z_l(x)$, corresponding respectively to $j_l(x)$, $h_l^{(1)}(x)$, $h_l^{(2)}(x)$. The constant $f(0)$ gives the appropriate normalization.

We work out one case in detail and leave the rest as exercises. Consider the case

$$Z_l(x) = f(0)x^l \int_{-i}^{i} e^{xu}\,(u^2+1)^l\,du \; . \tag{19.6.81}$$

Since this function is finite at the origin and is a solution of (19.6.69) it must be proportional to $j_l(x)$. To make it equal to $j_l(x)$ we choose $f(0)$ appropriately. Now, for $x \to 0$ we have

$$j_l(x) \to \frac{x^l}{(2l+1)!!} \quad \text{for} \quad x \to 0 \; . \tag{19.6.82}$$

But we also have that

$$Z_l(x) \to f(0)x^l \int_{-i}^{i} (u^2+1)^l\,du \equiv f(0)x^l I_l \quad \text{for} \quad x \to 0 \tag{19.6.83}$$

where we have made the identification

$$I_l = \int_{-i}^{i} (u^2+1)^l\,du \; . \tag{19.6.84}$$

To evaluate this integral put $u = it$ then

$$I_l = i\int_{-1}^{1}(1-t^2)^l\,dt\,. \tag{19.6.85}$$

Integrating by parts one can now show (problem 19.1) that

$$I_l = \frac{2l}{2l+1}\,I_{l-1} \quad l \geq 1\,. \tag{19.6.86}$$

Since $I_0 = 2i$, we immediately get

$$I_l = \frac{2^l l!}{(2l+1)!!}i\,. \tag{19.6.87}$$

Thus, we have

$$Z_l(x) \to if(0)\frac{2^{l+1}l!}{(2l+1)!!}x^l \quad \text{for} \quad x \to 0\,. \tag{19.6.88}$$

But as stated earlier

$$j_l(x) \to \frac{x^l}{(2l+1)!!} \quad \text{for} \quad x \to 0\,. \tag{19.6.89}$$

So, if we choose

$$if(0) = \frac{1}{2^{l+1}l!} \tag{19.6.90}$$

we get

$$j_l(x) = \frac{x^l}{i2^{l+1}l!}\int_{-i}^{i} e^{xu}\,(u^2+1)^l\,du\,. \tag{19.6.91}$$

Letting $u = it$ again this can be rewritten as

$$j_l(x) = \frac{x^l}{2^{l+1}l!}\int_{-1}^{1}(1-t^2)^l\cos xt\,dt\,. \tag{19.6.92}$$

The various properties of $j_l(x)$ can now be easily derived from this integral representation (problem 19.2).

In a similar manner we define

$$h_l^{(1)}(x) = \frac{x^l}{i2^l l!}\int_{\infty e^{i\alpha}}^{i} e^{xu}\,(u^2+1)^l\,du \tag{19.6.93}$$

$$h_l^{(2)}(x) = \frac{x^l}{i2^l l!}\int_{-i}^{\infty e^{i\alpha}} e^{xu}\,(u^2+1)^l\,du \tag{19.6.94}$$

where $\pi/2 \leq \alpha \leq \pi$ and $0 \leq \arg x \leq \pi/2$. For completeness we now list some of the most useful properties of the spherical Bessel functions. If one takes the integral representations as definitions of the spherical Bessel functions then all their properties can be derived from them.

Asymptotic Behaviour

$$j_l(x) \to \frac{1}{x}\cos[x-(l+1)\pi/2] = \frac{1}{x}\sin(x-l\pi/2) \quad \text{for} \quad x\to\infty \tag{19.6.95}$$

$$n_l(x) \to \frac{1}{x}\sin[x-(l+1)\pi/2] = -\frac{1}{x}\cos(x-l\pi/2) \quad \text{for} \quad x\to\infty \tag{19.6.96}$$

$$h_l^{(1)}(x) \to \frac{1}{x}\, e^{i[x-(l+1)\pi/2]} \quad \text{for} \quad x\to\infty \tag{19.6.97}$$

$$h_l^{(2)}(x) \to \frac{1}{x}\, e^{-i[x-(l+1)\pi/2]} \quad \text{for} \quad x\to\infty \ . \tag{19.6.98}$$

This shows that $h_l^{(1)}(x)$ corresponds asymptotically to an outgoing spherical wave and $h_l^{(2)}(x)$ corresponds asymptotically to an incoming spherical wave.

Behaviour for Small Argument

$$j_l(x) \to \frac{x^l}{(2l+1)!!} \quad \text{for} \quad x\to 0 \tag{19.6.99}$$

$$n_l(x) \to -\frac{(2l-1)!!}{x^{l+1}} \quad \text{for} \quad x\to 0 \ . \tag{19.6.100}$$

Also since $h_l^{(1)}(x) = j_l(x) + in_l(x)$ and $h_l^{(2)}(x) = j_l(x) - in_l(x)$ their behaviour for small values of x is given by

$$h_l^{(1)}(x) \to -i\frac{(2l-1)!!}{x^{l+1}} \quad \text{for} \quad x\to 0 \tag{19.6.101}$$

$$h_l^{(2)}(x) \to i\frac{(2l-1)!!}{x^{l+1}} \quad \text{for} \quad x\to 0 \ . \tag{19.6.102}$$

This shows that $j_l(x)$ is the only one of these functions that is finite at $x=0$.

We are now in a position to derive the partial wave decomposition of a plane wave, a result that we will need in the immediate future. Thus, we write

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{l,m} a_l(kr) Y^*_{l,m}(\hat{k}) Y_{l,m}(\hat{r}) \tag{19.6.103}$$

where as before $\hat{k}$, $\hat{r}$ indicate the directions of the corresponding variables.

If we let θ be the angle between $\mathbf{k}$ and $\mathbf{r}$ so that $\cos\theta = \hat{k}\cdot\hat{r}$ and then use the addition theorem for the spherical harmonics, namely

$$\sum_{m=-l}^{l} Y^*_{l,m}(\hat{k}) Y_{l,m}(\hat{r}) = \frac{2l+1}{4\pi} P_l(\cos\theta) \tag{19.6.104}$$

the expansion (19.6.103) can be written

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} a_l(kr) P_l(\cos\theta) \ . \tag{19.6.105}$$

Now using the orthogonality relation for the Legendre polynomials and calling $\cos\theta = u$ we get

$$a_l(kr) = 2\pi \int_{-1}^{1} e^{ikru}\, P_l(u)\, du\,. \tag{19.6.106}$$

But from the Rodrigues formula for the Legendre polynomials (equation (9.5.178) we have

$$P_l(u) = \frac{(-1)^l}{2^l l!}\frac{d^l}{du^l}(1-u^2)^l\,. \tag{19.6.107}$$

Substituting this into (19.6.106) and integrating by parts l times we get

$$a_l(kr) = 4\pi\, i^l \frac{(kr)^l}{2^{l+1}l!}\int_{-1}^{1} e^{ikru}\,(1-u^2)^l\, du\,. \tag{19.6.108}$$

Comparing this with (19.6.92) gives us the desired formula

$$e^{ikz} = \sum_{l=0}^{\infty}(2l+1)i^l j_l(kr)P_l(\cos\theta)\,. \tag{19.6.109}$$

19.7 Partial Wave Analysis

We now apply the tools developed in the previous section to a study of scattering from a spherically symmetric potential $V(r)$. Throughout this section we assume that $V(r)$ is short range in the sense that

$$\int r^2 V(r)\, dr < \infty\,. \tag{19.7.110}$$

Not only does this ensure that we have a well-defined scattering problem but it also ensures that various expressions, that we shall encounter, converge.

Consider the Schrödinger equation for a particle in such a spherically symmetric potential written in the form (19.2.3)

$$\left(\nabla^2 + k^2 - U(r)\right)\psi(\mathbf{r}) = 0\,. \tag{19.7.111}$$

As before we now decompose $\psi(\mathbf{r})$ into spherical harmonics

$$\psi(\mathbf{r}) = 4\pi\sum_{l,m} i^l R_l(r) Y^*_{lm}(\hat{k})Y_{lm}(\hat{r})\,. \tag{19.7.112}$$

The factor $4\pi i^l$ is included for later convenience. In terms of $\cos\theta = \hat{k}\cdot\hat{r}$ this expression becomes

$$\psi(\mathbf{r}) = \sum_{l=0}^{\infty} i^l(2l+1)R_l(r)P_l(\cos\theta)\,. \tag{19.7.113}$$

In either case substituting into (19.7.111) yields the radial equation for $R_l(r)$

$$\left[\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}-\frac{l(l+1)}{r^2}+k^2-U(r)\right]R_l(r)=0\ . \tag{19.7.114}$$

We still need to specify the boundary conditions on $R_l(r)$. For $r\to 0$ we want $R_l(r)$ finite. Thus, if $U(r)$ is bounded near the origin by c/r for some constant c then (19.7.114) reduces for $r\to 0$ to

$$\left[\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}-\frac{l(l+1)}{r^2}\right]R_l(r)=0 \tag{19.7.115}$$

and the solution finite near the origin is r^l. Thus,

$$R_l(r)\to A_l r^l \quad \text{for} \quad r\to 0\ . \tag{19.7.116}$$

The remaining boundary condition is determined by the physics of the situation that far from the potential we have the incoming beam plus the scattered beam. We start with an incoming beam $e^{i\mathbf{k}\cdot\mathbf{r}}$ which can be decomposed in each partial wave into an incoming spherical wave and an outgoing spherical wave as follows:

$$\begin{aligned}
e^{i\mathbf{k}\cdot\mathbf{r}} &= \sum_{l=0}^{\infty} i^l(2l+1)j_l(kr)P_l(\cos\theta)\\
&= \frac{1}{2}\sum_{l=0}^{\infty} i^l(2l+1)\left[h_l^{(1)}(kr)+h_l^{(2)}(kr)\right]P_l(\cos\theta)
\end{aligned} \tag{19.7.117}$$

where we have used equations (19.6.67) and (19.6.68).

The effect of the potential is to cause scattering and hence modify the amplitude of each of the *outgoing* spherical waves. Thus, asymptotically the full solution must be of the form

$$\begin{aligned}
\psi(\mathbf{r}) &\to \sum_{l=0}^{\infty} i^l(2l+1)\frac{1}{2}\left[h_l^{(2)}(kr)+S_l(k)h_l^{(1)}(kr)\right]P_l(\cos\theta)\\
&\quad\ \text{for} \ r\to\infty
\end{aligned} \tag{19.7.118}$$

where S_l contains the total effect of the scattering. We have incorporated here the fact that for $r\to\infty$, $\psi(\mathbf{r})$ must satisfy the free Schrödinger equation due to the short-range nature of the potential as well as the fact that *the scattering process affects only the outgoing waves*. Equation (19.7.118) can now be rewritten as

$$\begin{aligned}
&\psi(\mathbf{r})\to\\
&\frac{1}{2}\sum_{l=0}^{\infty} i^l(2l+1)\left[h_l^{(1)}(kr)+h_l^{(2)}(kr)+[S_l(k)-1]h_l^{(1)}(kr)\right]P_l(\cos\theta)\\
&= \sum_{l=0}^{\infty} i^l(2l+1)\left[j_l(kr)+\frac{1}{2}[S_l(k)-1]h_l^{(1)}(kr)\right]P_l(\cos\theta)
\end{aligned}$$

$$= \quad e^{i\mathbf{k}\cdot\mathbf{r}} + \sum_{l=0}^{\infty} i^l (2l+1)\frac{1}{2}\left[S_l(k) - 1\right] h_l^{(1)}(kr) P_l(\cos\theta)$$

$$\text{for } r \to \infty \,. \tag{19.7.119}$$

Since this is the solution for large r we can replace $h_l^{(1)}(kr)$ by its asymptotic form (19.6.97) to get the final asymptotic form for $\psi(\mathbf{r})$.

$$\psi(\mathbf{r}) \to e^{i\mathbf{k}\cdot\mathbf{r}} + f(k,\theta)\frac{e^{ikr}}{r} \tag{19.7.120}$$

where

$$f(k,\theta) = \frac{-i}{2k}\sum_{l=0}^{\infty}(2l+1)\left[S_l(k) - 1\right] P_l(\cos\theta) \ . \tag{19.7.121}$$

The asymptotic expression (19.7.120) is precisely of the form we previously derived quite generally (equations (19.2.4), (19.2.8)) and thus shows that $f(k,\theta)\, e^{ikr}/$ is the scattered wave. Thus, $f(k,\theta)$ can again be identified as the scattering amplitude. Recalling (19.2.10) we see that the differential cross-section $d\sigma/d\Omega$ is given by

$$\frac{d\sigma}{d\Omega} = |f(k,\theta)|^2 \ . \tag{19.7.122}$$

Thus, our scattering problem is solved if we find $f(k,\theta)$ or alternatively $S_l(k)$. The remaining boundary condition for $R_l(r)$ can now be extracted from these results.

$$R_l(r) \to \frac{1}{2}\left[h_l^{(2)}(kr) + S_l(k) h_l^{(1)}(kr)\right] \quad \text{for } \ r \to \infty \tag{19.7.123}$$

or, using the asymptotic forms of the spherical Hankel functions we obtain

$$R_l(r) \to \frac{1}{2kr}\left[e^{-i[kr-(l+1)\pi/2]} + S_l(k) e^{i[kr-(l+1)\pi/2]}\right] \quad \text{for } \ r \to \infty \,. \tag{19.7.124}$$

19.8 Phase Shifts

For elastic scattering, from a spherical potential, probability must be conserved for each partial wave since different l values do not couple (L^2 commutes with H). Thus, the magnitude of the radial flux, through a sphere of radius r, for each incoming and outgoing partial wave must be the same but in opposite directions. The incoming partial waves are given by $h_l^{(2)}(kr)$. The outgoing partial waves are given by $S_l(k) h_l^{(1)}(kr)$. In both cases we can compute the flux, through a very large sphere, using their asymptotic form. Equating the magnitude of the two fluxes yields

$$-\lim_{r\to\infty} 4\pi r^2 \frac{\hbar}{2im}\left[\frac{e^{i[kr-(l+1)\pi/2]}}{kr}\frac{d}{dr}\frac{e^{-i[kr-(l+1)\pi/2]}}{kr}\right.$$

$$
\begin{aligned}
& -\frac{e^{-i[kr-(l+1)\pi/2]}}{kr}\frac{d}{dr}\frac{e^{i[kr-(l+1)\pi/2]}}{kr}\Bigg] \\
= & \lim_{r\to\infty} 4\pi r^2 \frac{\hbar}{2im}|S_l(k)|^2 \left[\frac{e^{-i[kr-(l+1)\pi/2]}}{kr}\frac{d}{dr}\frac{e^{i[kr-(l+1)\pi/2]}}{kr}\right. \\
& \left. -\frac{e^{i[kr-(l+1)\pi/2]}}{kr}\frac{d}{dr}\frac{e^{-i[kr-(l+1)\pi/2]}}{kr}\right] . \qquad (19.8.125)
\end{aligned}
$$

Writing this out we find

$$|S_l(k)|^2 = 1 \ . \qquad (19.8.126)$$

Thus, for elastic scattering, the S-matrix elements, $S_l(k)$ may be expressed in terms of real numbers $\delta_l(k)$ called the *phase shifts* and defined by

$$S_l(k) = e^{2i\delta_l(k)} \qquad (19.8.127)$$

where the factor of two is conventional.

These phase shifts have an intuitive interpretation which we now describe. For this purpose we consider a potential of finite range so that

$$U(r) = 0 \quad \text{for} \quad r \geq a \ . \qquad (19.8.128)$$

Then, for $r \geq a$ the most general solution of the radial equation is

$$R_l(r) = A_l\left[\cos\delta_l\, j_l(kr) - \sin\delta_l\, n_l(kr)\right] \quad r \geq a \ . \qquad (19.8.129)$$

The asymptotic behaviour of this solution for large r is

$$R_l(r) \to \frac{A_l}{kr}\left[\cos\delta_l \sin(kr - l\pi/2) + \sin\delta_l \cos(kr - l\pi/2)\right] \qquad (19.8.130)$$

or

$$R_l(r) \to \frac{A_l}{kr}\sin(kr - l\pi/2 + \delta_l) \quad \text{for} \quad r \to \infty \ . \qquad (19.8.131)$$

On the other hand the free incoming wave is

$$j_l(r) \to \frac{A_l}{kr}\sin(kr - l\pi/2) \quad \text{for} \quad r \to \infty \ . \qquad (19.8.132)$$

Thus, the effect of the potential is to shift the phase of the solution in the region of no interaction. In a later section we demonstrate this result explicitly for a square well potential.

The phase shifts here defined coincide with the ones in equation (19.8.127). This also agrees with the definition of phase shifts given for the one-dimensional scattering problems in section (5.8). If we now use the definitions of phase shifts to compute the differential cross-section we obtain

$$
\begin{aligned}
& \frac{d\sigma}{d\Omega} = |f(k,\theta)|^2 \\
= & \frac{1}{k^2}\sum_{l,l'=0}^{\infty}(2l+1)(2l'+1)\sin\delta_l \sin\delta_{l'}\cos(\delta_l - \delta_{l'}) \times \\
& \times P_l(\cos\theta)P_{l'}(\cos\theta) \ . \qquad (19.8.133)
\end{aligned}
$$

Here we have used the fact that the sums over l and l' are symmetric in l and l' to replace $e^{i(\delta_l-\delta_{l'})}$ by the symmetrized term

$$\frac{1}{2}\left[e^{i(\delta_l-\delta_{l'})}+e^{i(\delta_{l'}-\delta_l)}\right]=\cos(\delta_l-\delta_{l'})\;. \qquad (19.8.134)$$

In general these are infinite sums and difficult to evaluate. However, as we now indicate, for potentials of short or finite range and for low energies, these sums may be truncated after just a few terms. In other words, only a few partial waves contribute.

Suppose that for $r \geq a$ the potential may be neglected. Then for this region the total energy of a particle is kinetic $(\hbar\mathbf{k})^2/2m$ and exceeds $[\hbar^2 l(l+1)]/(2mr^2)$, the energy due to orbital motion alone

$$\frac{(\hbar\mathbf{k})^2}{2m}\geq\frac{\hbar^2 l(l+1)}{2ma^2} \qquad (19.8.135)$$

or

$$k^2a^2\geq l(l+1) \qquad (19.8.136)$$

whence we get the inequality

$$l\leq ka\;. \qquad (19.8.137)$$

Thus, if l violates this inequality the corresponding partial wave cannot participate in the scattering classically and we may truncate the sums in (19.8.133) by the largest $l < ka$.

Clearly for potentials of finite range the length a is well defined. For short range potentials, however, it is not as clear how to perform the approximation. In section 19.11 we replace the above heuristic argument with a rigorous derivation and prove that those higher partial waves that violate (19.8.137) do not contribute appreciably to the differential cross-section.

19.9 The Optical Theorem: Unitarity Bound

If we integrate (19.8.133) to get the total cross-section we find using

$$\int P_l(\cos\theta)P_{l'}(\cos\theta)\,d\Omega=\frac{4\pi}{2l+1}\delta_{l,l'} \qquad (19.9.138)$$

that all the cross terms disappear and

$$\sigma=\frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l\;. \qquad (19.9.139)$$

On the other hand

$$\begin{aligned}\Im(f(k,\theta)) &= \frac{1}{2k}\sum_{l=0}^{\infty}(2l+1)\Re(1-S_l)P_l(\cos\theta)\\ &= \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l P_l(\cos\theta)\;. \qquad (19.9.140)\end{aligned}$$

If we now set $\theta = 0$ and use $P_l(1) = 1$ we get

$$\Im(f(k,0)) = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l \ . \tag{19.9.141}$$

Comparing this with (19.9.139) we obtain the optical theorem

$$\sigma = \frac{4\pi}{k}\Im(f(k,0)) \, . \tag{19.9.142}$$

This relation is very useful for experimentalists since it allows them to get a good approximation for the total cross section by measuring the scattering amplitude as close as possible to the forward ($\theta = 0$) direction. This relation also shows why it is so difficult to measure the total cross section; it is necessary to measure scattering in the forward direction, which is precisely the direction that the unscattered beam goes.

The total cross-section (19.9.139) can also be written

$$\sigma = \sum_{l=0}^{\infty}\sigma_l \tag{19.9.143}$$

where the lth partial cross-section σ_l is given by

$$\sigma_l = \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l \, . \tag{19.9.144}$$

For elastic scattering, the phase shifts are real and it therefore follows that for elastic scattering

$$\sigma_l \leq \frac{4\pi}{k^2}(2l+1) \, . \tag{19.9.145}$$

The value $(4\pi/k^2)(2l+1)$ is known as the *unitarity bound* and is reached only for $\delta_l = (n+1/2)\pi$. This is the condition for *resonance* and shows up as a local maximum in the cross-section for the corresponding partial wave.

19.10 Partial Waves: Lippmann-Schwinger Equation

The Lippmann-Schwinger equation in configuration space reads

$$\psi^{(\pm)}(\mathbf{r}) = \psi^{(0)}(\mathbf{r}) + \int G^{\pm}(\mathbf{r},\mathbf{r}')V(r')\psi^{(\pm)}(\mathbf{r}')\,d^3r' \tag{19.10.146}$$

where $G^{\pm}(\mathbf{r},\mathbf{r}')$ are the free Green's functions satisfying

$$(\nabla^2 + k^2)G^{\pm}(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') \tag{19.10.147}$$

with outgoing and incoming wave boundary conditions.

To obtain the partial wave decompositions of these equations, we set

$$\begin{aligned}\psi^{(\pm)}(\mathbf{r}) &= \sum_{l,m}\psi_l^{(\pm)}(r)Y^*_{lm}(\hat{k})Y_{lm}(\hat{r})\\ \psi^{(0)}(\mathbf{r}) &= \sum_{l,m}\psi_l^{(0)}(r)Y^*_{lm}(\hat{k})Y_{lm}(\hat{r})\end{aligned}\tag{19.10.148}$$

where as always $\hat{k}$, $\hat{r}$ indicate the θ, φ directions of $\mathbf{k}$ and $\mathbf{r}$ respectively. We also expand the Green's functions

$$G^{\pm}(\mathbf{r},\mathbf{r}') = \sum_{lm} G_l^{(\pm)}(r,r')Y^*_{lm}(\hat{r}')Y_{lm}(\hat{r}) \ . \tag{19.10.149}$$

We then substitute all of these expressions into (19.10.146) and integrate out explicitly the angular variables corresponding to $\mathbf{r}'$. This yields the expression

$$\psi_l^{(\pm)}(r) = \psi_l^{(0)}(r) + \int G_l^{(\pm)}(r,r')V(r')\psi_l^{(\pm)}(r')\,r'^2dr' \tag{19.10.150}$$

where we have also used the fact that the potential is spherically symmetric and the spherical harmonics form a complete orthonormal set. The explicit form for the partial wave Green's functions $G_l^{(\pm)}(r,r')$ is obtained by substituting the expansion (19.10.149) into the defining equation (19.10.147) for the full Green's functions $G^{\pm}(\mathbf{r},\mathbf{r}')$. This yields

$$\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} + k^2\right) G_l^{(\pm)}(r,r') = \frac{1}{rr'}\delta(r-r') \tag{19.10.151}$$

where the partial wave Green's functions are finite at $r = 0$ and have asymptotic behaviour specified by the superscripts $(\pm)$. These Green's functions are furthermore continuous at $r = r'$ and have a discontinuity in their first derivative. This discontinuity is obtained by integrating both sides of equation (19.10.151) about $r = r'$ to get

$$\frac{d}{dr}G_l^{(\pm)}(r,r')\bigg|_{r=r'+0} - \frac{d}{dr}G_l^{(\pm)}(r,r')\bigg|_{r=r'-0} = \frac{1}{r'^2} \ . \tag{19.10.152}$$

We therefore have, to begin with, as general solutions of (19.10.151)

$$\begin{aligned}&G_l^{(\pm)}(r,r') = A_l^{(\pm)}(r')\,j_l(kr) \quad && r < r'\\ &G_l^{(+)}(r,r') = B_l^{(+)}(r')\,h_l^1(kr) && r > r'\\ &G_l^{(-)}(r,r') = B_l^{(-)}(r')\,h_l^2(kr) && r > r' \ .\end{aligned}\tag{19.10.153}$$

We now impose the boundary conditions: continuity of these functions as well as the given discontinuity in the first derivatives to get the "constants" $A_l^{(\pm)}(r')$, $B_l^{(\pm)}(r')$.

$$\begin{aligned}A_l^{(+)}(r') &= -ikh_l^{(1)}(kr')\\ A_l^{(-)}(r') &= ikh_l^2(kr')\\ B_l^{(\pm)}(r') &= \mp ikj_l(kr') \ .\end{aligned}\tag{19.10.154}$$

Thus, the partial wave Green's functions are given by

$$\begin{aligned} G_l^{(+)}(r,r') &= -ikj_l(kr_<)h_l^{(1)}(kr_>) && (19.10.155)\\ G_l^{(-)}(r,r') &= ikj_l(kr_<)h_l^{(2)}(kr_>) && (19.10.156) \end{aligned}$$

where

$$r_< = \begin{cases} r & \text{for} \quad r < r' \\ r' & \text{for} \quad r > r' \end{cases}$$

$$r_> = \begin{cases} r & \text{for} \quad r > r' \\ r' & \text{for} \quad r < r' \end{cases} . \qquad (19.10.157)$$

In obtaining these results we have also used the Wronskian (see problem 19.4)

$$j_l(x)\,n_l'(x) - n_l(x)\,j_l'(x) = \frac{1}{x^2} \,. \qquad (19.10.158)$$

This completes the partial wave decomposition of the Lippmann-Schwinger equations. They are given by (19.10.150) with the explicit forms (19.10.155) and (19.10.156) for $G^{\pm}(\mathbf{r},\mathbf{r}')$ respectively. These equations are extremely useful, since they are completely equivalent to the radial Schrödinger equation plus boundary conditions. It is this fact, that the integral equations already incorporate the boundary conditions, that makes them so useful.

19.11 Effective Range Approximation

It is frequently desirable to have a parametrization of scattering data in terms of variables more readily interpreted than phase shifts. For low energy scattering a set of such variables is provided by the scattering length and effective range. In this section we develop expressions for these two parameters. We furthermore supplant the heuristic argument of section 19.8, that only a few partial waves contribute for low energy scattering, by rigourous bounds.

To this end, we consider the Schrödinger equation for the lth partial wave in a spherically symmetric potential

$$\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} + k^2 - U(r)\right) R_l(r) = 0 \,. \qquad (19.11.159)$$

We want to compare the solution to this equation with the solution for the lth partial wave of the free Schrödinger equation that is finite at the origin

$$\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} + k^2\right) j_l(kr) = 0 \,. \qquad (19.11.160)$$

To compare these solutions we multiply (19.11.159) by $j_l(kr)$ and (19.11.160) by $R_l(r)$ and subtract to get

$$\begin{aligned} & j_l(kr)\frac{d^2R_l(r)}{dr^2} - R_l(r)\frac{d^2j_l(kr)}{dr^2} + \frac{2}{r}\left[j_l(kr)\frac{dR_l(r)}{dr} - R_l(r)\frac{dj_l(kr)}{dr}\right] \\ & = j_l(kr)U(r)R_l(r) \,. \qquad (19.11.161) \end{aligned}$$

We next integrate from 0 to some point a to be fixed later, and use integration by parts on the first term to get

$$\begin{aligned}
&\int_0^a r^2 dr \frac{d}{dr}\left[j_l(kr)\frac{dR_l(r)}{dr} - R_l(r)\frac{dj_l(kr)}{dr}\right] \\
+ \; & 2\int_0^a r dr \left[j_l(kr)\frac{dR_l(r)}{dr} - R_l(r)\frac{dj_l(kr)}{dr}\right] \\
= \; & a^2 \left[j_l(kr)\frac{dR_l(r)}{dr} - R_l(r)\frac{dj_l(kr)}{dr}\right]\Bigg|_{r=a} \\
= \; & \int_0^a r^2 dr j_l(kr) U(r) R_l(r) \,. \qquad (19.11.162)
\end{aligned}$$

Now we choose a sufficiently large value for a so that $R_l(a)$ as well as $j_l(ka)$ assume their asymptotic forms

$$\begin{aligned}
R_l &\approx e^{i\delta_l}\left[\cos\delta_l \, j_l(ka) - \sin\delta_l \, n_l(ka)\right] \\
&\approx e^{i\delta_l}\left[\cos\delta_l \, \frac{\sin(ka - l\pi/2)}{ka} + \sin\delta_l \, \frac{\cos(ka - l\pi/2)}{ka}\right] \qquad (19.11.163)
\end{aligned}$$

and we already used

$$j_l(ka) \approx \frac{\sin(ka - l\pi/2)}{ka} \quad , \quad n_l(ka) \approx -\frac{\cos(ka - l\pi/2)}{ka} \quad . \qquad (19.11.164)$$

So,

$$R_l(a) \approx e^{i\delta_l} \, \frac{\sin(ka - l\pi/2 + \delta_l)}{ka} \,. \qquad (19.11.165)$$

Substituting these results into (19.11.161) we obtain the following integral representation for the phase shifts

$$e^{i\delta_l} \sin\delta_l \approx k \int_0^a dr \, r^2 j_l(kr) \, U(r) \, R_l(r) \,. \qquad (19.11.166)$$

For a sufficiently large value of a this equation becomes exact. To obtain an estimate for the phase shift, we replace $R_l(r)$ by $\exp(i\delta_l) \, j_l(kr)$. This extremely crude approximation assumes δ_l is small and yields

$$\sin\delta_l \approx k \int_0^a dr \, r^2 j_l^2(kr) \, U(r) \,. \qquad (19.11.167)$$

A much more accurate approach using a variational technique was introduced by Schwinger [19.4] and used by Blatt and Jackson [19.5]. Now suppose the potential has a range ρ and the energy is sufficiently low so that $k\rho << 1$. In this case we can (firstly) replace the upper limit in the integral by ρ and (secondly) use the asymptotic behaviour for small argument of the Bessel functions. This yields the relation

$$\sin\delta_l \approx \frac{(k\rho)^{2l+1}}{[(2l+1)!!]^2}\rho \int_0^\rho dr \, U(r) \left(\frac{r}{\rho}\right)^{2(l+1)} \quad . \qquad (19.11.168)$$

From this relation we can make the following deductions:

1) The phase shifts are odd functions of k.

2) For low energies ($k\rho << 1$) the phase shifts decrease rapidly with l.

In order to prove this second observation we maximize the integral by setting $r = \rho$. Also we replace $\sin\delta_l$ by δ_l. Then we get

$$\frac{\delta_1}{\delta_0} = \left(\frac{k\rho}{3}\right)^3 \quad , \quad \frac{\delta_2}{\delta_0} = \frac{(k\rho)^4}{(15)^2} \quad , \quad \text{etc.}$$

This shows that for low energy the scattering occurs predominantly in the lower partial waves, particularly the $l = 0$ or s-waves. This also justifies our previous heuristic arguments that partial wave analysis is useful for low energy scattering since only a few partial waves contribute. With this in mind we now examine in more detail the s-wave scattering for low energy. Restricting the discussion to s-wave solutions, we can repeat the previous computation starting from the radial equations

$$\begin{aligned} \left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} + k^2 - U(r)\right] R_0(k,r) &= 0 \\ \left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - U(r)\right] R_0(0,r) &= 0\,. \end{aligned} \tag{19.11.169}$$

Again, multiplying the first of these by by $R_0(0,r)$, and the second by $R_0(k,r)$ and subtracting we obtain

$$\begin{aligned} & R_0(0,r)R_0''(k,r) - R_0(k,r)R_0''(0,r) \\ + & \frac{2}{r}\left[R_0(0,r)R_0'(k,r) - R_0(k,r)R_0'(0,r)\right] \\ = & -k^2 R_0(0,r)R_0(k,r)\,. \end{aligned} \tag{19.11.170}$$

For large r we have the asymptotic form of $R_0(k,r)$

$$R_0(k,r) \approx c\,\frac{\sin(kr+\delta_0)}{r} \equiv f(k,r)\,. \tag{19.11.171}$$

This function $f(k,r)$ clearly satisfies the free $l = 0$ radial Schrödinger equation as does $f(0,r)$. Hence we obtain

$$\begin{aligned} & f(0,r)f''(k,r) - f(k,r)f''(0,r) \\ + & \frac{2}{r}\left[f(0,r)f'(k,r) - f(k,r)f'(0,r)\right] \\ = & -k^2 f(0,r)f(k,r)\,. \end{aligned} \tag{19.11.172}$$

If we further choose

$$c = \frac{1}{\sin\delta_0} \tag{19.11.173}$$

then

$$\begin{aligned} f(0,r) &= \lim_{k\to 0}\frac{1}{\sin\delta_0}\left[\frac{\sin\delta_0\cos kr}{r}+\frac{\cos\delta_0\sin kr}{r}\right] \\ &= \frac{1}{r}+k\cot\delta_0 \\ &= \frac{1}{r}\left(1-\frac{r}{c_0}\right) \end{aligned} \tag{19.11.174}$$

where

$$\frac{1}{c_0} = -\lim_{k\to 0} k\cot\delta_0\ . \tag{19.11.175}$$

We now subtract (19.11.172) from (19.11.170) and integrate the resultant equation from 0 to ∞. If we further use the easily computed result that

$$r^2\left[f(k,r)f'(0,r)-f(0,r)f'(k,r)\right]_{r=0} = \frac{1}{c_0}+k\cot\delta_0 \tag{19.11.176}$$

we get

$$k\cot\delta_0 = -\frac{1}{c_0}+k^2\int_0^\infty r^2dr\ \left[f(k,r)f(0,r)-R_0(k,r)R_0(0,r)\right]\ . \tag{19.11.177}$$

This result is usually referred to as the *Bethe formula.* It is useful for short range potentials such as are encountered in nuclear physics when one can divide space into two regions:

1) an internal region $r < r_0$, $kr_0 < 1$ where the potential $|V| > E$ and

2) an external region $r > r_0$ where the potential $|V| < E$.

To obtain the effective range approximation we make one final low energy approximation and set $k = 0$ in the integral and define

$$r_0 = 2\int_0^\infty r^2dr\left[f^2(0,r)-R_0^2(0,r)\right]\ . \tag{19.11.178}$$

The final formula for the low energy s-wave phase shift in the effective range approximation is then

$$k\ \cot\delta_0 \approx -\frac{1}{c_0}+\frac{1}{2}k^2r_0\ . \tag{19.11.179}$$

The constant c_0 is known as the *scattering length* and r_0 is called the *effective range.* Actually equation (19.11.179) says only that $k\ \cot\delta_0$ is an even function of k for small k. The usefulness of this formula derives from the interpretation that can be given to it for potentials strong compared to the energy. For smooth potentials that do not change sign, r_0 is proportional to the range of the potentials. However for more complicated potentials r_0 may even be negative and the above simple interpretation does not follow. Since for "simple" potentials r_0 depends only on the range and depth of the potential, equation (19.11.179) is also known as the *shape-independent* approximation.

The scattering length is related to the existence of bound states. From (19.11.173) we see that c_0 gives the location of the zero in $f(0,r)$, the zero-energy form of the asymptotic solution (19.11.171). If the potential has a range ρ and a depth V_0 then for $V_0\rho$ so small that there are no s-wave bound states we have that

$$\frac{\pi}{2} > \delta_0 > 0$$

and, as we see from (19.11.175), $c_0 < 0$.

If $V_0\rho$ is at the transitional strength for producing the first bound state,

$$\delta_0 = \frac{\pi}{2}$$

and $c_0 = \pm\infty$.

For $V_0\rho$ sufficiently strong to produce a bound state,

$$\frac{\pi}{2} < \delta_0 < \pi$$

so that $c_0 > 0$. The three cases are depicted in figs.19.1, 19.2, and 19.3 with a corresponding square well potential.

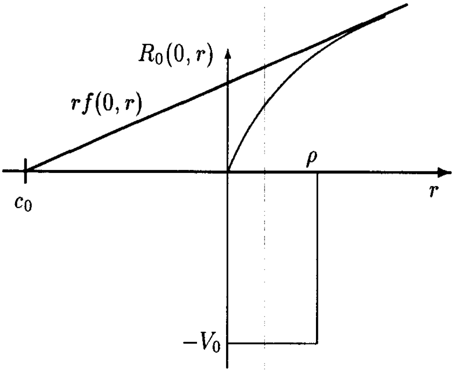

Figure 19.1: No s-wave bound states: $\pi/2 > \delta_0 > 0\,,\ c_0 < 0$.

19.12 Resonant Scattering

One of the simplest potentials for which the scattering problem can be solved in closed form is the square well. We use this potential to illustrate some of the results to be expected in potential scattering, in general. Also there are many instances in which more realistic potentials can be approximated by an appropriate square well potential. For a square well we have with $V_0 > 0$

$$V(r) = \left\{ \begin{array}{cc} -V_0 & r \le a \\ 0 & r > a \end{array} \right. . \qquad (19.12.180)$$

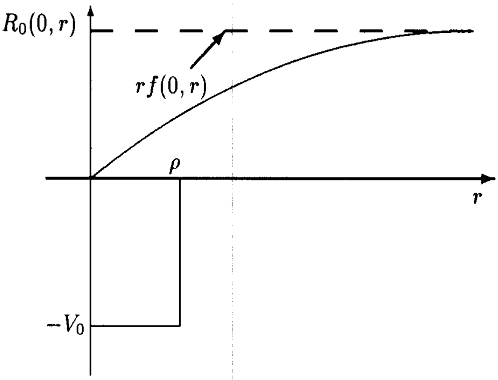

Figure 19.2: Transitional strength : $\delta_0 = \pi/2$, $c_0 = \pm\infty$.

If we now define, as usual,

$$k^2 = \frac{2mE}{\hbar^2} \quad , \quad K^2 = \frac{2m(E+V_0)}{\hbar^2} \tag{19.12.181}$$

the radial equation becomes

$$\begin{aligned}\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} + K^2\right] R_l(r) &= 0 \quad r \le a \\ \left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} + k^2\right] R_l(r) &= 0 \quad r > a \ .\end{aligned} \tag{19.12.182}$$

The solutions of these equations satisfying the conditions that $R_l(0)$ is finite and $R_l(r)$ corresponds to a fixed incoming flux for large values of r are

$$R_l(r) = \begin{cases} A_l j_l(Kr) & r \le a \\ j_l(kr) + \frac{1}{2}\left[S_l(k) - 1\right] h_l^{(1)}(kr) & r > a \end{cases} \ . \tag{19.12.183}$$

Replacing the $S_l(k)$ by the phase shifts we obtain an alternate expression for $r > a$.

$$R_l(r) = e^{i\delta_l}\left[\cos\delta_l j_l(kr) - \sin\delta_l n_l(kr)\right] \tag{19.12.184}$$

where we have also used the relation $h_l^{(1)} = j_l + in_l$. Since R_l and its first derivatives are continuous we can match the solutions at $r = a$. To eliminate the irrelevant constant A_l it is convenient to match logarithmic derivatives. We thus get

$$\frac{Kj_l'(Ka)}{j_l(Ka)} = \frac{k\left[\cos\delta_l j_l(ka) - \sin\delta_l n_l(ka)\right]}{\cos\delta_l j_l(ka) - \sin\delta_l n_l(ka)} \tag{19.12.185}$$

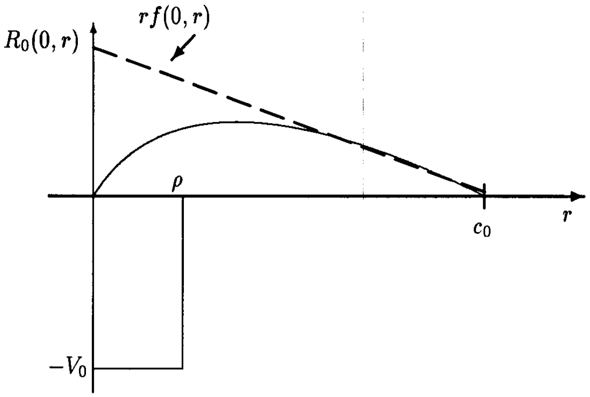

Figure 19.3: One s-wave bound state: $\pi/2 < \delta_0 < \pi$, $c_0 > 0$.

where the primes indicate derivatives with respect to the argument. A little algebra then yields

$$\tan\delta_l = \frac{kj_l(Ka)j_l'(ka) - Kj_l'(Ka)j_l(ka)}{kj_l(Ka)n_l'(ka) - Kj_l'(Ka)n_l(ka)} \ . \tag{19.12.186}$$

To continue with a more general discussion we consider an arbitrary smooth potential of finite range a. Then for $r > a$, that is, outside the range of the potential, the solution is as before. We can now write a formal expression for the phase shift in terms of the logarithmic derivative

$$\frac{1}{\gamma_l(k)} = \frac{a}{R_l}\frac{dR_l}{dr}\bigg|_{r=a} \ . \tag{19.12.187}$$

Then,

$$\tan\delta_l = \frac{ka\,\gamma_l(k)j_l'(ka) - j_l(ka)}{ka\,\gamma_l(k)n_l'(ka) - n_l(ka)} \ . \tag{19.12.188}$$

However, in the general case $\gamma_l(k)$ is not given by

$$\frac{1}{\gamma_l(k)} = \frac{Ka\,j_l'(Ka)}{j_l(Ka)} \tag{19.12.189}$$

but must be found by solving the radial equation for $r \leq a$.

For low energies we can use the asymptotic form for the spherical Bessel functions and find

$$\tan\delta_l(k) = \frac{(2l+1)(ka)^{2l+1}}{[(2l+1)!!]^2}\,\frac{l\gamma_l(k) - 1}{(l+1)\gamma_l(k)+1} \ . \tag{19.12.190}$$

Now, in general, for $ka \to 0$ we see that $\delta_l(k)$ tends to zero. However, if it happens that the potential is such that for some value $E_0 = (\hbar^2 k_0^2)/2m$ of the energy we have that $(l+1)\gamma_l(k_0)+1 \to 0$ then $\delta_l(k_0)$ can approach $\pi/2$. This is resonance. In this case the phase shift may be approximated near E_0 by

$$\tan\delta_l(k) = \frac{\Gamma/2}{E_0 - E} \; . \tag{19.12.191}$$

The corresponding partial cross-section is

$$\begin{aligned}\sigma_l &= \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l(k) = \frac{4\pi}{k^2}(2l+1)\frac{1}{1+\cot^2\delta_l} \\ &= \frac{4\pi}{k^2}(2l+1)\frac{\Gamma^2/4}{(E_0-E)^2+\Gamma^2/4} \; . \end{aligned} \tag{19.12.192}$$

This is known as a *Breit-Wigner* form and clearly displays the resonance at E_0 with a width Γ. The cross-section reaches its maximum allowed value, the *unitarity bound*

$$\frac{4\pi}{k^2}(2l+1)$$

at the resonance. This generalizes our discussion of resonances that we started with alpha-decay in Chapter 14.

19.13 Problems

19.1 Derive equation (19.6.86)

19.2 Use the integral representation for the spherical Bessel functions $Z_l(x)$ to show that

a)

$$Z_{l-1}(x) + Z_{l+1}(x) = \frac{2l+1}{x} Z_l(x) \quad \text{for} \quad l \geq 1$$

b)

$$\frac{d}{dx} Z_l(x) = Z_{l-1}(x) - \frac{l+1}{x} Z_l(x) \quad \text{for} \quad l \geq 1$$

c)

$$\frac{d}{dx}\left[x^{-l} Z_l(x)\right] = -x^{-l} Z_{l+1}(x) \; .$$

Here $Z_l(x)$ may be any one of the four spherical Bessel functions.

19.3 A useful formula for generating any one of the spherical Bessel functions is the Rodrigues formula

$$Z_l(x) = x^l \left(-\frac{1}{x}\frac{d}{dx}\right)^l Z_0 \; .$$

Use the results of problem 19.2 to get

$$Z_l(x) = -x^{l-1}\frac{d}{dx}\left(\frac{Z_{l-1}}{x^{l-1}}\right)$$

and hence derive the generating formula above.

19.4 Show that the Wronskian

$$W = j_l(x)n_l'(x) - n_l(x)j_l'(x)$$

satisfies the differential equation

$$\frac{dW}{dx} = -\frac{2}{x}W \ .$$

Solve this equation and use the behaviour of $j_l(x)$, $n_l(x)$ for small x to fix the constant of integration to get

$$W = \frac{1}{x^2} \ .$$

Hint: Start with the differential equations for $j_l(x)$ and $n_l(x)$.

19.5 Find the equation for the phase shifts for scattering off a hard sphere potential.

$$V(r) = \begin{cases} \infty & \text{for} \quad r = a \\ 0 & \text{for} \quad r > a \end{cases} \ .$$

Solve this for $l = 0, 1$.

19.6 a) Calculate the differential cross section in first Born approximation for the potential

$$V(r) = V_0\, e^{-\mu r^2} \ .$$

b) To the same approximation compute the s-wave ($l = 0$) phase shift.

19.7 Repeat problem 19.6 for the potential

$$V(r) = \begin{cases} -V_0 & \text{for} \quad r < a \\ 0 & \text{for} \quad r \geq a \end{cases} \ .$$

19.8 Compute the phase shifts for scattering by a potential

$$V(r) = V_0 a\, \delta(r - a) \ .$$

19.9 Compute approximate $l = 0$ and $l = 1$ phase shifts for scattering a high energy particle of mass m by a short range potential

$$V(r) = V_0 \frac{e^{-\alpha r}}{r} \ .$$

Use whatever seems to be an appropriate approximation.

19.10 Consider a potential "shell" of value V_0 between $r = a$ and $r = b$ and zero otherwise. Calculate the s-wave phase shift and show that for large V_0 (with respect to what?) resonances occur approximately at energies which would be bound states if the particles were confined to a "box" between $r = a$ and $r = b$.

19.11 Verify directly by using the differential equation and their behaviour near $x = 0$ that

$$j_0(x) = \frac{\sin x}{x} \quad \text{and} \quad n_0(x) = -\frac{\cos x}{x} \ .$$

19.12 Given the potential

$$V(r) = \begin{cases} -V_0 & r < a \\ 0 & r > a \end{cases}$$

find the effective range and the scattering length for the s-wave ($l = 0$).

19.13 Repeat problem 19.12 for the Yukawa potential

$$V = -V_0 \frac{e^{-\mu r}}{r} \ .$$

19.14 Use the results of problems 19.12 and 19.13 to fix the parameters of the Yukawa potential in terms of those of the square well so that both yield the same s-wave scattering length and effective range. The fact that this is possible is what is meant by calling this a "shape-independent" approximation.

19.15 a) Find the phase shifts for scattering by a hard sphere

$$V(r) = \begin{cases} \infty & r < a \\ 0 & r > a \end{cases} \ .$$

b) Find the total cross-section for an incoming energy

$$E = \frac{\hbar^2 k^2}{2m}$$

in the two limits:

$$k \to 0$$

$k \to \infty$.

Give a physical explanation for the factors of 4 and 2.
Hint: For $k \to \infty$ use the asymptotic forms of j_l and n_l to obtain a simple form for $\sin^2 \delta_l$. Furthermore, replace the sum over l by an integral so that

$$\sigma = \sum_{l=0}^{l=ka} \sigma_l \approx \frac{4\pi}{k^2} \int_0^{ka} (2l+1) \sin^2 \delta_l \, dl \, .$$

19.16 Use the exact solution for the square well to find the condition on the potential for the s-wave ($l = 0$) to produce a resonance at an energy $E_0 = (\hbar^2 k_0^2)/2m$.

19.17 a) Find the differential cross-section for scattering of 3.0 eV electrons from a double slit. Assume the slits are cut into very thin material and that the incident beam is normal to the plane of the slits. Also assume that the scattering is weak.
b) Calculate the minimum intensity of the incident beam so that if a detector is 2.5 m from the slits at the position of the first maximum, beyond the central maximum, and is the width of the first maximum it receives a counting rate of 200 electrons/s.
c) Is this first maximum resolved from the central maximum according to the Rayleigh criterion?

19.18 Show that for a spherically symmetric potential $V(r)$ the total cross-section in first Born approximation is given by

$$\sigma = \frac{m^2}{2\pi\hbar^4 k^2} \int_0^{2k} |\tilde{V}(q)|^2 \, q dq$$

where $\tilde{V}$ is the Fourier transform of V. Use this result together with the properties of the Fourier transform to conclude that for high energies and a potential of finite range a the scattering is appreciable only in the forward direction where the scattering angle θ satisfies

$$\sin(\theta/2) \leq \frac{1}{2ka} \, .$$

Bibliography

[19.1] A very readable little book on integral equations is:
F. Smithies, *Integral Equations* - Cambridge University Press (1962).
All the references of chapter 18 are applicable to this chapter as well.

[19.2] An enjoyable book treating the more mathematical aspects of potential scattering is:
V. de Alfaro and T. Regge, *Potential Scattering* - North-Holland Pub. Co., Amsterdam (1965).

[19.3] The Born approximation was first derived and applied to quantum mechanical scattering problems in:
M. Born, Z. Physik, **37**, 863, 1926. It was in this paper that Born also first introduced the probability interpretation for the wavefunction.

[19.4] J. Schwinger, Phys. Rev. **72**, 742, (1947).

[19.5] J.M. Blatt and J.D. Jackson, Phys. Rev. **76**, 18, (1949).
see also
H. Bethe, Phys. Rev. **76**, 38, (1949).

Chapter 20

Systems of Identical Particles

20.1 Introduction

One of the most profound and far-reaching consequences of quantum mechanics results from the indistinguishability of two identical particles. That two identical particles are indistinguishable seems a tautology. Nevertheless classically it is possible to follow (at least in principle) the trajectory of any particle. Thus, in classical mechanics, if two identical particles interact we can, in principle, follow each particle even throughout the region of interaction until they are separated. In this sense the particles retain their individuality and are in fact distinguishable, in principle. This means that in classical mechanics, no real difference occurs in the treatment of a system of distinguishable or indistinguishable particles.

In quantum mechanics the situation is very different and that is why the discussion has been delayed until now. If we consider two identical particles that come together, interact and then separate, their individuality is lost. This occurs because during the interaction their wavefunctions must overlap (occupy the same portion of space at the same time). The wavefunction is no longer a product of two functions, but just a single function which depends on the variables for both particles. When the particles move apart it is impossible to tell which particle was which. There is simply one complicated function describing both particles. This is a consequence of the fact that we cannot follow the individual trajectories of the two particles and since they are identical we cannot say, even when they are again separated, where each one came from.

Of course this is not built into the theory a priori and we must now do just that. To simplify the discussion we begin by considering a system of only two identical particles. Later we generalize this to an arbitrary number.

20.2 Two Identical Particles

Consider a system of two identical particles interacting with each other via a two-body potential $V_{int}(\mathbf{x}_1, \mathbf{x}_2)$. Since, according to Newton's third law, action and reaction are equal and opposite we must have

$$V_{int}(\mathbf{x}_1, \mathbf{x}_2) = V_{int}(\mathbf{x}_2, \mathbf{x}_1) \ . \tag{20.2.1}$$

In addition to the two-body potential we can have an external force, described by a potential $V(\mathbf{x})$, acting on the particles. The Hamiltonian for the two particles is then given by

$$H = T_1 + T_2 + V(\mathbf{x}_1) + V(\mathbf{x}_2) + V_{int}(\mathbf{x}_1, \mathbf{x}_2) \tag{20.2.2}$$

where T_i is the kinetic energy for the particle labelled i. Thus

$$T_i = \frac{\mathbf{p}_i^2}{2m} \ . \tag{20.2.3}$$

We write this Hamiltonian as $H(1,2)$ to indicate the dependence on the particle labels. The indistinguishability of the particles is reflected in the fact that

$$H(1,2) = H(2,1) \ . \tag{20.2.4}$$

If we let $\psi(1,2)$ be the wavefunction for the two particles then the Schrödinger equation reads

$$H(1,2)\psi(1,2) = E\,\psi(1,2) \ . \tag{20.2.5}$$

We now define a particle-exchange operator P_{12} with the property that

$$P_{12}\psi(1,2) = \psi(2,1) \ . \tag{20.2.6}$$

Then as a consequence of (20.2.4) we have that

$$P_{12}H(1,2)P_{12}^{-1} = H(2,1) = H(1,2) \tag{20.2.7}$$

so that

$$[P_{12}, H(1,2)] = 0 \ . \tag{20.2.8}$$

It therefore follows that the eigenstates of $H(1,2)$ can be labelled with the eigenvalues E of $H(1,2)$ and the eigenvalues α of P_{12} as well as whatever other labels are necessary. The eigenvalues of P_{12} are easily found since as (20.2.6) shows

$$P_{12}^2 = 1 \ . \tag{20.2.9}$$

This implies that

$$\alpha^2 = 1 \quad , \quad \alpha = \pm 1 \ . \tag{20.2.10}$$

The two eigenvalues correspond to two physically very different types of particles, known as *bosons* for $\alpha = +1$ and *fermions* for $\alpha = -1$.

For bosons the wavefunction is *symmetric* under the interchange of particle labels

$$\psi(1,2) = \psi(2,1) \ . \tag{20.2.11}$$

For fermions the wavefunction is *antisymmetric* under the interchange of a pair of particle labels

$$\psi(1,2) = -\psi(2,1) \ . \tag{20.2.12}$$

These simple rules when generalized to a system of N particles have very far-reaching consequences. Before carrying out this generalization we consider in more detail one specific problem involving two electrons.

20.3 The Hydrogen Molecule

The hydrogen molecule consists of two hydrogen atoms bound together. Thus, we have two protons and two electrons as shown in fig. 20.1. Since the protons are much more massive than the electrons we neglect their motion and treat them as fixed centres of force. With this approximation the Hamiltonian for a hydrogen molecule becomes

$$H = \frac{\mathbf{p}_1^2}{2m} - \frac{e^2}{r_1} + \frac{\mathbf{p}_2^2}{2m} - \frac{e^2}{r_2} + \frac{e^2}{r_{12}} + \frac{e^2}{R} - \frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}} \tag{20.3.13}$$

where the various quantities are labelled in fig. 20.1. We write this Hamiltonian as

$$H = H_0 + H' \tag{20.3.14}$$

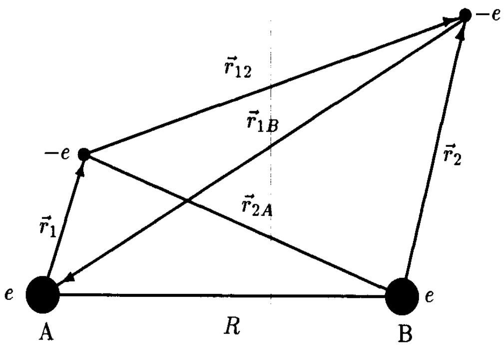

Figure 20.1: Geometry of the hydrogen molecule.

where H_0 is the "hydrogen atom" part of the Hamiltonian

$$H_0 = \frac{\mathbf{p}_1^2}{2m} - \frac{e^2}{r_1} + \frac{\mathbf{p}_2^2}{2m} - \frac{e^2}{r_2} \tag{20.3.15}$$

and

$$H' = \frac{e^2}{r_{12}} + \frac{e^2}{R} - \frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}} \ . \tag{20.3.16}$$

The Hamiltonian (20.3.13) is of precisely the same form as the Hamiltonian (20.2.2). Here the two-body potential is e^2/r_{12}, whereas $-(e^2/r_1 + e^2/r_{1B})$ is the "external" potential for particle 1 and $-(e^2/r_2 + e^2/r_{2A})$ is the "external" potential for particle 2. The repulsive potential e^2/R between the two protons is just a constant.

To find the ground state energy of the hydrogen molecule we treat H' as a perturbation. In that case we only need to find a good approximation for the ground state wavefunction. The energy is then approximated by the expectation value of the total Hamiltonian, H in this approximate ground state. Now for R sufficiently large we have two unperturbed hydrogen atoms with the Hamiltonian H_0. We use this for our approximate ground state.

Electrons have spin 1/2 and therefore are fermions. Thus, the *total wavefunction must be antisymmetric* in all the particle labels, and hence we must include the spin. There are two states of good total spin:
$s = 1$, the triplet state χ^t, which is even in the two particle labels, and
$s = 0$, the singlet state χ^s, which is odd in the two particle labels.
The corresponding spatial wavefunctions are therefore,

$$\psi_s(\mathbf{r}_1, \mathbf{r}_2) = A_s \left[\psi_A(1)\psi_B(2) + \psi_A(2)\psi_B(1)\right] \tag{20.3.17}$$

and

$$\psi_t(\mathbf{r}_1, \mathbf{r}_2) = A_t \left[\psi_A(1)\psi_B(2) - \psi_A(2)\psi_B(1)\right] \tag{20.3.18}$$

where A_s and A_t are normalization constants, and $\psi_A(j)$, $\psi_B(k)$ represent hydrogen atom ground state wave functions of electron j and k centred at A and B respectively. Thus,

$$\begin{aligned} \psi_A(1) &= \frac{1}{\sqrt{\pi a^3}}\, e^{-r_1/a} \quad , \quad & \psi_B(2) &= \frac{1}{\sqrt{\pi a^3}}\, e^{-r_2/a} \\ \psi_A(2) &= \frac{1}{\sqrt{\pi a^3}}\, e^{-r_{2A}/a} \quad , \quad & \psi_B(1) &= \frac{1}{\sqrt{\pi a^3}}\, e^{-r_{1B}/a} \end{aligned} \tag{20.3.19}$$

where a is the Bohr radius appropriate for a nucleus of chage $2e$

$$a = \frac{\hbar^2}{2me^2} \ . \tag{20.3.20}$$

The normalizations A_t and A_s are given by

$$1 = |A_s|^2 \left(1 + 2|(\psi_A, \psi_B)|^2 + 1\right) \tag{20.3.21}$$

$$1 = |A_s|^2 \left(1 - 2|(\psi_A, \psi_B)|^2 + 1\right) \tag{20.3.22}$$

so

$$A_s = \frac{1}{\sqrt{2(1+A^2)}} \tag{20.3.23}$$

$$A_t = \frac{1}{\sqrt{2(1-A^2)}} \tag{20.3.24}$$

where the overlap integral A is given by

$$A = (\psi_A, \psi_B) = \frac{1}{\pi a^3} \int e^{-(r_1 + r_{1B})/a} d^3 r_1 \ . \tag{20.3.25}$$

To evaluate this integral we transform to elliptical coordinates

$$\xi = \frac{r_1 + r_{1B}}{R} \quad , \quad \eta = \frac{r_1 - r_{1B}}{R} \quad , \quad \varphi \tag{20.3.26}$$

where φ is the angle of rotation about the line joining the two protons. The volume element in these coordinates is [20.3]

$$d^3 r_1 = \frac{R^3}{8} (\xi^2 - \eta^2) d\xi d\eta d\varphi \tag{20.3.27}$$

with the range of integration

$$1 \leq \xi < \infty \quad -1 \leq \eta \leq 1 \quad 0 \leq \varphi \leq 2\pi \ . \tag{20.3.28}$$

Thus,

$$\begin{aligned} A &= \frac{1}{8\pi} \left(\frac{R}{a}\right)^3 \int_1^\infty e^{-(R\xi)/a}\, d\xi \int_{-1}^{1} (\xi^2 - \eta^2)\, d\eta \int_0^{2\pi} d\varphi \\ &= \left[1 + \frac{R}{a} + \frac{1}{3} + \left(\frac{R}{a}\right)^2\right] e^{-R/a} \ . \end{aligned} \tag{20.3.29}$$

To the extent that ψ_s and ψ_t are good wavefunctions, the ground state energy of the hydrogen molecule can be approximated in the singlet and triplet states by

$$\begin{aligned} E_s(R) &= \int \psi_s(1,2)\, H \psi_s(1,2)\, d^3 r_1\, d^3 r_2 \\ E_t(R) &= \int \psi_t(1,2)\, H \psi_t(1,2)\, d^3 r_1\, d^3 r_2 \ . \end{aligned} \tag{20.3.30}$$

These integrals can be rewritten with a little algebra in the form

$$\begin{aligned} E_s(R) - 2E_0 &= \frac{J + K}{1 + A^2} \\ E_t(R) - 2E_0 &= \frac{J - K}{1 - A^2} \end{aligned} \tag{20.3.31}$$

where

$$E_0 = -\frac{e^2}{2a} = \text{the ground state energy of the hydrogen atom} \tag{20.3.32}$$

$$\begin{aligned} J &= \int |\psi_A(1)\psi_B(2)|^2 \left(\frac{e^2}{r_{12}} - \frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}} + \frac{e^2}{R}\right) d^3r_1\, d^3r_2 \\ &= \frac{e^2}{R} - \int |\psi_A(1)|^2 \frac{e^2}{r_{1B}}\, d^3r_1 - \int |\psi_B(2)|^2 \frac{e^2}{r_{2A}}\, d^3r_2 \\ &+ \int |\psi_A(1)|^2 \frac{e^2}{r_{12}} |\psi_B(2)|^2\, d^3r_1\, d^3r_2 \end{aligned} \tag{20.3.33}$$

and

$$\begin{aligned} K &= \int \psi_A(1)\psi_B(2) \left(\frac{e^2}{r_{12}} - \frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}} + \frac{e^2}{R}\right) \psi_A(2)\psi_B(1)\, d^3r_1\, d^3r_2 \\ &= \frac{e^2}{R}|A|^2 - A\int \psi_B(2)\frac{e^2}{r_{2A}}\psi_A(2)\, d^3r_2 - A\int \psi_A(1)\frac{e^2}{r_{1B}}\psi_B(1)\, d^3r_1 \\ &+ \int \psi_A(1)\psi_B(2)\frac{e^2}{r_{12}}\psi_A(2)\psi_B(1)\, d^3r_1\, d^3r_2 \ . \end{aligned} \tag{20.3.34}$$

The integral J is called the *Coulomb Interaction Integral* and contains the following terms.

1. The first term gives the Coulomb repulsion of the two protons.
2. The second term gives the energy due to the interaction of the proton at B with the "charge density" $-e|\psi_A(1)|^2$ due to electron 1 at A.
3. The third term gives the interaction energy between the proton at A and the "charge density" $-e|\psi_B(2)|^2$ of electron 2 at B and is therefore equal to the second term.
4. The last term gives the interaction between the two "charge densities" $-e|\psi_A(1)|^2$ and $-e|\psi_B(2)|^2$ of the two electrons centred at A and B.

The integral K is something totally foreign to classical physics. It is called the *exchange energy* and results strictly from the indistinguishability of the two electrons. It is this type of term that gives rise to *covalent bonding.* The various integrals are evaluated as follows. The two integrals

$$\int |\psi_A(1)|^2 \frac{e^2}{r_{1B}}\, d^3r_1 = \int |\psi_B(2)|^2 \frac{e^2}{r_{2A}}\, d^3r_2 \tag{20.3.35}$$

are evaluated by the use of elliptical coordinates (eqns. (20.3.26) - (20.3.28)) and yield

$$\int |\psi_A(1)|^2 \frac{e^2}{r_{1B}}\, d^3r_1 = \frac{e^2R^2}{2a^3}\left[\int_1^\infty \xi\, e^{-R\xi/a}\, d\xi \int_{-1}^{1} e^{-R\eta/a}\, d\eta\right.$$

$$\begin{aligned} &+ \left.\int_1^\infty e^{-R\xi/a}\,d\xi \int_{-1}^{1} \eta\, e^{-R\eta/a}\,d\eta\right] \\ &= \frac{e^2}{R}\left[1 - e^{-2R/a}\left(1+\frac{R}{a}\right)\right] . \end{aligned} \tag{20.3.36}$$

Also we can write

$$\int |\psi_A(1)|^2 \frac{e^2}{r_{12}} |\psi_B(2)|^2\, d^3r_1\, d^3r_2 = \int \rho_A(1)\phi_B(1)\, d^3r_1 \tag{20.3.37}$$

where

$$\rho_A(1) = -e|\psi_A(1)|^2 \tag{20.3.38}$$

is the "charge density" due to electron 1 centred at A and $\phi_B(1)$ is the "potential" at r_1 due to electron 2 with charge density $-e|\psi_B(2)|^2$ centred at B. Thus $\phi_B(1)$ satisfies the Poisson equation

$$\nabla^2\phi_B(1) = 4\pi e|\psi_B(2)|^2 \ . \tag{20.3.39}$$

This "potential" is calculated in several books on electrostatics [20.1]. Hence we can combine all these results and get

$$J = \frac{e^2}{R} e^{-2R/a}\left[1 + \frac{5}{8}\frac{R}{a} - \frac{3}{4}\left(\frac{R}{a}\right)^2 - \frac{1}{6}\left(\frac{R}{a}\right)^3\right] . \tag{20.3.40}$$

For the exchange integral K the terms involving only one integration are again obtained by going to elliptical coordinates. The last term cannot be expressed in terms of elementary functions. It can, however, be expressed in terms of exponential integrals

$$E_i(x) = -\int_x^\infty \frac{e^{-t}}{t}\,dt \ . \tag{20.3.41}$$

This was done by Sugiura [20.2]. The final result for K is

$$\begin{aligned} K &= \frac{e^2}{R}A^2\left[1 + \frac{6}{5}\left(C + \ln\frac{R}{a}\right)\right] \\ &+ \frac{e^2}{a}e^{-2R/a}\left[\frac{11}{8} + \frac{103}{20}\frac{R}{a} + \frac{49}{15}\left(\frac{R}{a}\right)^2 + \frac{11}{15}\left(\frac{R}{a}\right)^3\right] \\ &+ \frac{e^2}{R}e^{-2R/a}\left[1 - \frac{R}{a} + \frac{1}{3}\left(\frac{R}{a}\right)^2\right] E_i(-4R/a) \\ &- \frac{12}{5}A\,e^{R/a}\left[1 - \frac{R}{a} + \frac{1}{3}\left(\frac{R}{a}\right)^2\right] E_i(-2R/a) \end{aligned} \tag{20.3.42}$$

where $C = 0.577215\ldots$ is Euler's constant.

If all these results are combined we obtain the results sketched in fig. 20.2 for the energies $E_s(R)$, $E_t(R)$.

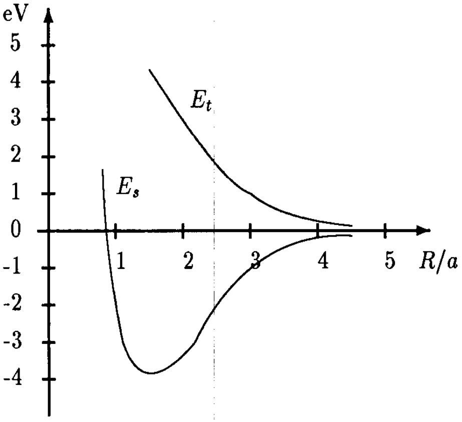

Figure 20.2: Energy as a function of separation.

Although the quantitative agreement with experiment is rather poor for the above computation, the qualitative features are correct and are shown in fig. 20.2. The computation can again be improved by replacing ψ_A, ψ_B by wavefunctions with an effective proton charge Ze and treating Z as a variational parameter. This improves the quantitative agreement with experiment considerably. The main point of this calculation, however, was to bring out the effect of the exchange energy. This energy is a purely quantum mechanical effect and is entirely a consequence of the antisymmetry of the electron wavefunction under the interchange of the two electrons. We next consider what happens in the case of several particles.

20.4 *N* Identical Particles

We now take up again the discussion started in section 2, except that we do not limit ourselves to two particles. The argument follows very closely the argument of section 2. The general Hamiltonian for a system of N identical particles takes the form

$$H(1,2,\ldots,N) = \sum_{i=1}^{N} \left[\frac{p_i^2}{2m} + V(i)\right] + V_{int}(1,2,\ldots,N) \tag{20.4.43}$$

where $V(i)$ is an external potential acting on the ith particle and $V_{int}(1,2,\ldots,N)$ represents the mutual interaction between the particles. If, as is usually the case, the interaction between the particles is due only to two-body interactions then

$$V_{int}(1,2,\ldots,N) = \frac{1}{2}\sum_{i\neq j}^{N} V_{int}(i,j) \tag{20.4.44}$$

where again due to Newton's third law

$$V_{int}(i,j) = V_{int}(j,i) \ . \tag{20.4.45}$$

The labels i, j etc. represent all of the particle coordinates, *internal as well as external.* Also the factor of 1/2 in front is to avoid double counting (counting the same interaction twice). Combining equations (20.4.43) - (20.4.45) we see that the N-particle Hamiltonian is invariant under the interchange of particles. Thus, it is again possible to introduce a particle exchange operator P_{ij} that commutes with the total Hamiltonian. The action of P_{ij} is

$$P_{ij}\psi(1,2,\ldots,i,\ldots,j,\ldots,N) = \psi(1,2,\ldots,j,\ldots,i,\ldots,N) \ . \tag{20.4.46}$$

As for the case of two particles the eigenvalues of P_{ij} are ±1 and can be used together with the energy E to label the eigenstates of the Hamiltonian.

It is a remarkable fact of nature, that greatly simplifies all considerations, that for a given type of particle all the eigenvalues of P_{ij} are either $+1$ or -1 and are never mixed. Particles with the eigenvalue $+1$ are called bosons and particles with the eigenvalue -1 are called fermions.

To illustrate these considerations we consider a wavefunction for three identical particles and write out the $(3! = 6)$ symmetry conditions.

1) Three identical bosons

$$\psi(1,2,3) = \psi(1,3,2) = \psi(2,1,3) = o\psi(2,3,1) = \psi(3,1,2) = \psi(3,2,1) \ .$$

2) Three identical fermions

$$\psi(1,2,3) = -\psi(1,3,2) = -\psi(2,1,3) = \psi(2,3,1) = \psi(3,1,2) = -\psi(3,2,1) \ .$$

The nature of a given type of particle is furthermore intimately connected with its spin. This result is summarized in the following theorem.

Spin and Statistics Theorem

All particles with integer spin (including spin zero) are bosons and all particles with half-odd integer spins are fermions.

The proof of this theorem is beyond the scope of this book since it requires the machinery of relativistic quantum field theory. We, therefore, content ourselves with the statement of this theorem. Experimentally it has been verified to a very high degree for electrons and also for photons.

There is one result, the *Pauli Exclusion Principle*, which follows quite straightforwardly from the requirement that a fermion wavefunction is totally antisymmetric. For a system of N identical fermions no two particles may occupy the same state.

This result follows quite trivially. Let $\psi_{k_1 \ldots k_i \ldots k_j \ldots k_N}(x_1 \ldots x_i \ldots x_j \ldots x_N)$ be the totally antisymmetric wave function for a system of N non-interacting identical fermions. Here k_i is a complete labelling of the state occupied by particle i. If such a labelling is possible also for interacting fermions then the rest of our argument also applies to them. The total antisymmetry means that

$$\begin{aligned} & \psi_{k_1\ldots k_i\ldots k_j\ldots k_N}(x_1\ldots x_i\ldots x_j\ldots x_N) \\ = & -\psi_{k_1\ldots k_i\ldots k_j\ldots k_N}(x_1\ldots x_j\ldots x_i\ldots x_N) \\ = & -\psi_{k_1\ldots k_j\ldots k_i\ldots k_N}(x_1\ldots x_i\ldots x_j\ldots x_N) \ . \end{aligned} \tag{20.4.47}$$

If we now assume that two particles are in the same state, or what is equivalent that $k_i = k_j$ then we have the further result that

$$\begin{aligned} & \psi_{k_1\ldots k_i\ldots k_j\ldots k_N}(x_1\ldots x_i\ldots x_j\ldots x_N) \\ = & \ \psi_{k_1\ldots k_j\ldots k_i\ldots k_N}(x_1\ldots x_i\ldots x_j\ldots x_N) \ \text{ for } \ k_i = k_j \ . \end{aligned} \tag{20.4.48}$$

Combining this result with eqn. (20.4.47) we obtain that for identical fermions

$$\psi_{k_1\ldots k_i\ldots k_i\ldots k_N}(x_1\ldots x_i\ldots x_j\ldots x_N) = 0 \ . \tag{20.4.49}$$

This result is known as the *Pauli Exclusion Principle.* It states that no two fermions (both interacting as well as non-interacting) may simultaneously occupy the same state. Under state we must understand here *all* the quantum numbers of the particles involved. Thus, particles with different mass or charge automatically have different states. This principle was originally postulated by Pauli to provide an explanation for the periodic table. Its consequence are, however, of much greater generality. Thus, for example, many of the properties of solids are a consequence of this principle.

20.5 Non-Interacting Fermions

Consider a system of N fermions that do not interact with each other but are subject to an external potential $V(x)$. The Hamiltonian for this system is

$$H = \sum_{i=1}^{N}\left[\frac{p_i^2}{2m} + V(x_i)\right] = \sum_{i=1}^{N} H_i \ . \tag{20.5.50}$$

Since the index i simply labels the various particles, their corresponding single particle Hamiltonians H_i commute

$$[H_i, H_j] = 0 \tag{20.5.51}$$

and hence

$$[H_i, H] = 0 \ . \tag{20.5.52}$$

Thus, to diagonalize the N-particle Hamiltonian (20.5.50) we need only diagonalize the single-particle Hamiltonians. This is, of course, a consequence of the unrealistic assumption of no interaction between the particles.

We now let $|k\rangle$, $k = 0, 1, 2, \ldots$ be a complete set of one-particle states satisfying the one-particle Schrödinger equation

$$H_i|k\rangle = E_k|k\rangle\,. \tag{20.5.53}$$

An eigenstate of the total Hamiltonian is then given by $\prod_{i=1}^{N}|k_i\rangle$. This state must of course be totally antisymmetrized. If $\mathcal{A}$ is an antisymmetrization operator then the N particle fermion eigenfunction of the Hamiltonian (20.5.50) is

$$|k_1, k_2, \ldots, k_N\rangle = \mathcal{A}|k_1\rangle|k_2\rangle \ldots |k_N\rangle\,. \tag{20.5.54}$$

Here k_i is the value of k for the i-th particle. Clearly, we also have

$$H|k_1, k_2, \ldots, k_N\rangle = (E_{k_1} + E_{k_2} + \cdots + E_{k_N})|k_1, k_2, \ldots, k_N\rangle\,. \tag{20.5.55}$$

We shall now repeat the procedure above for a fixed representation. This will make it easy to carry out the appropriate antisymmetrization.

The one-particle states in configuration space are given by

$$\phi_k(x_j) = \langle x_j|k\rangle\,. \tag{20.5.56}$$

We again emphasize that the labels k are a *complete set of labels for the one-particle states* and that x_j denotes the configuration space variable for the j-th particle. In this fixed coordinate representation, a totally antisymmetric eigenstate of the full Hamiltonian is given by a so-called *Slater determinant.* Thus, if the single-particle wavefunctions are normalized then the antisymmetrized and normalized N-particle wave function is given by the following determinant known as a Slater determinant.

$$\Psi_{k_1\ldots k_N}(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}}\det\begin{vmatrix} \phi_{k_1}(x_1) & \phi_{k_1}(x_2) & \cdots & \phi_{k_1}(x_N) \\ \phi_{k_2}(x_1) & \phi_{k_2}(x_2) & \cdots & \phi_{k_2}(x_N) \\ \cdot & \cdot & \cdots & \cdot \\ \phi_{k_N}(x_1) & \phi_{k_N}(x_2) & \cdots & \phi_{k_N}(x_N) \end{vmatrix}\,. \tag{20.5.57}$$

The antisymmetry of a determinant under the interchange of rows or columns makes it obvious that this wavefunction has the required antisymmetry. The fact that the wavefunction can be written as a sum of products of single-particle wavefunctions is again a consequence of the unphysical assumption of no interaction between the individual particles.

Nevertheless a wave-function of the form of a Slater determinant is frequently the starting point for a computation involving an interacting fermion system. The main justification for this is the simplicity of the approach.

20.6 Non-Interacting Bosons

We now repeat the discussion of the previous section for a system of N bosons that do not interact with each other but are subject to an external potential

$V(x)$. As before we have an Hamiltonian consisting of a sum of single-particle Hamiltonians

$$H = \sum_{i=1}^{N} \left[\frac{p_i^2}{2m} + V(x_i) \right] = \sum_{i=1}^{N} H_i \ . \tag{20.6.58}$$

We again start with a complete set of one-particles states $\{|k\rangle \ , \ k = 0, 1, 2, \ldots\}$ that are eigenstates of the single particle Hamiltonian

$$H_i|k\rangle = E_k|k\rangle \ . \tag{20.6.59}$$

An eigenstate of the total Hamiltonian is then given by a symmetrized product of such single particle states. Thus, we have the N boson state

$$|k_1, k_2, \ldots, k_N\rangle = \frac{1}{\sqrt{N!}} \sum_P |k_1\rangle|k_2\rangle \ldots |k_N\rangle \tag{20.6.60}$$

where the sum is over all $N!$ permutations of the N particle labels. It is again straightforward to go to a fixed representation and construct the analogue of a Slater determinant. We start with a normalized set of one-particle wavefunctions in configuration space

$$\phi_k(x_j) = \langle x_j|k\rangle \ . \tag{20.6.61}$$

In terms of these an N particle boson state is given by the following *permanent*

$$\Psi_{k_1 \ldots k_N}(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}} \mathrm{Perm} \begin{vmatrix} \phi_{k_1}(x_1) & \phi_{k_1}(x_2) & \cdots & \phi_{k_1}(x_N) \\ \phi_{k_2}(x_1) & \phi_{k_2}(x_2) & \cdots & \phi_{k_2}(x_N) \\ \cdot & \cdot & \cdots & \cdot \\ \phi_{k_N}(x_1) & \phi_{k_N}(x_2) & \cdots & \phi_{k_N}(x_N) \end{vmatrix} \tag{20.6.62}$$

where "Perm" means a permanent, which is the same as a determinant without changes of sign.

Thus, again we have the N-particle state expressed as a sum of products of one-particle states. As for the case of fermions this is again a consequence of the lack of interaction between the particles. In the next section we refine this method of description even further. It is to be remembered, however, that the labels are single-particle labels. Thus, we do not have a convenient machinery for handling collective phenomena, such as particle correlations or phase transitions, starting from a basis of one-particle states.

20.7 N-Space: Second Quantization for Bosons

In this section we develop an elegant method for handling the symmetrization of boson states. In order to do this we first generalize from a system with a definite number N of particles to a system with an arbitrary number of particles.

We again start with a complete basis of one particle states $\{|k\rangle \ , \ k = 0, 1, 2, \ldots\}$, and corresponding wavefunctions $\phi_k(x_j)$. An N-boson state can now be described by stating that particle 1 is in the state $|k_1\rangle$, particle 2 is in the state $|k_2\rangle$ etc. and then symmetrizing to get the state $|k_1, \ldots, k_N\rangle$.

Another but completely equivalent way of stating this is to say that there are n_0 particles in the state $|0\rangle$, n_1 particles in the state $|1\rangle$ etc. up to n_∞. The reason this specification of an N boson state by the sequence of integers

$$|n_0, n_1, \ldots, n_\infty\rangle$$

is equivalent to giving the state $|k_1, \ldots, k_N\rangle$ explicitly is due to the fact that the state $|k_1, \ldots, k_N\rangle$ is totally symmetric. Thus, we can define a state of N bosons in *occupation number space* or N-space by an infinite sequence of non-negative integers $|n_0, n_1, \ldots, n_\infty\rangle$ such that

$$\sum_{k=0}^{\infty} n_k = N \ . \tag{20.7.63}$$

In this way of specifying states, the four particle states $|0, 2, 5, 5\rangle$ corresponding to a symmetric product of the states $|0\rangle$, $|2\rangle$,$|5\rangle$,$|5\rangle$ is given by

$$|1, 0, 1, 0, 0, 2, 0, 0, 0, \ldots, 0, \ldots\rangle.$$

This says that there is one particle in the ground state $k = 0$, one particle in the state $k = 2$, two particles in the state $k = 5$, and no particles in any other state. If the one-particle states are the eigenstates of a one-particle Hamiltonian $H(x_j)$ such that the N-particle Hamiltonian is

$$H = \sum_{j=1}^{N} H(x_j) \tag{20.7.64}$$

then the state $|n_0, n_1, \ldots, n_\infty\rangle$ is an eigenstate of H. The eigenvalue of H is given in terms of the single particle energies E_k

$$H(x_j)|k\rangle = E_k|k\rangle \ . \tag{20.7.65}$$

In fact,

$$H|n_0, n_1, n_1, \ldots\rangle = \left(\sum_{k=0}^{\infty} n_k E_k\right) |n_0, n_1, n_1, \ldots\rangle \ . \tag{20.7.66}$$

It therefore becomes convenient to introduce operators N_k such that

$$N_k|n_0, n_1, n_1, \ldots\rangle = n_k|n_0, n_1, n_1, \ldots\rangle \ . \tag{20.7.67}$$

Thus, the operators N_k "count" the number of particles in the state $|k\rangle$ and are appropriately named *number operators*. The Hamiltonian may now be written

$$H = \sum_{k=0}^{\infty} N_k E_k \ . \tag{20.7.68}$$

We now take the further step and write N_k in terms of annihilation and creation operators like we did for the harmonic oscillator

$$N_k = a_k^\dagger a_k \tag{20.7.69}$$

where we assume the commutation relation

$$[a_k, a_j^\dagger] = \delta_{kj} \tag{20.7.70}$$

with all other operators commuting

$$[a_k, a_j] = [a_k^\dagger, a_j^\dagger] = 0 \ . \tag{20.7.71}$$

With this notation we now have

$$H = \sum_{k=0}^{\infty} E_k a_k^\dagger a_k \ . \tag{20.7.72}$$

Notice that a_k and $a_k^\dagger$ respectively annihilate and create quanta of energy E_k. The ground state (zero-particle state or vacuum) is $|0,0,0,\ldots,0,\ldots\rangle$ and is annihilated by all the annihilation operators. A general state can now be written as

$$|n_0, n_1, \ldots\rangle = \frac{(a_0^\dagger)^{n_0}}{\sqrt{n_0!}} \frac{(a_1^\dagger)^{n_1}}{\sqrt{n_1!}} \cdots |0,0,0,\ldots\rangle \ . \tag{20.7.73}$$

An N-particle state in configuration space is furthermore given by

$$\Psi_{n_0,n_1,\ldots}(x_1, x_2, \ldots, x_N) = \langle x_1, x_2, \ldots, x_N | n_0, n_1, \ldots\rangle \ . \tag{20.7.74}$$

Thus, the completeness relation for N-particle states in occupation number space is

$$\sum_{\{n\}} |n_0, n_1, \ldots\rangle\langle n_0, n_1, \ldots| \tag{20.7.75}$$

where the sum extends over all sequences $\{n_0, n_1, \ldots\}$ with a fixed number N of particles

$$\sum n_k = N \ . \tag{20.7.76}$$

It is therefore also convenient to define a *total number operator* N such that the eigenvalue of N for an N-particle state is N. This operator is given by

$$\mathsf{N} = \sum N_k = \sum a_k^\dagger a_k \ . \tag{20.7.77}$$

The orthonormality of these states in occupation number space is expressed by

$$\langle n_0, n_1, \ldots, n_k, \ldots | n_0', n_1', \ldots, n_k', \ldots\rangle = \prod_{k=0}^{\infty} \delta_{n_k, n_k'} \ . \tag{20.7.78}$$

It is therefore straightforward to write an arbitrary state (a state with an arbitrary number of particles) in terms of the occupation number space basis

$$|\Psi\rangle = \sum |n_0, n_1, \ldots\rangle\langle n_0, n_1, \ldots|\Psi\rangle \, . \tag{20.7.79}$$

The coefficients $\langle n_0, n_1, \ldots|\Psi\rangle$ are the probability amplitudes for finding n_0 particles in the state $|0\rangle$, n_1 particles in the state $|1\rangle$, etc. if the state of the system is $|\Psi\rangle$. For a fixed number N of particles these coefficients are, in fact, nothing else but the permanent given by eqn. (20.6.62).

In arriving at eqn. (20.7.73) we have in fact rewritten a system of N non-ineracting bosons like an infinite system of harmonic oscillators. This procedure, at present, is nothing other than an elegant formalism for incorporating the symmetry of bose states. It turns out, however, to have far reaching consequences in solid state physics where many computations are facilitated, and much physical insight is gained from this machinery.

20.8 N-Space: Second Quantization for Fermions

We now carry out a discussion parallel to that of the previous section except that we deal with a system of fermions. It is again convenient to generalize from a system with a fixed number N of particles to a system with an arbitrary number of particles.

As before let $\{|k\rangle \, , \; k = 0, 1, 2, \ldots\}$ be a complete basis of one-particle states. An N-fermion state is then given if we state that particle 1 is in the state $|k_1\rangle$, particle 2 is in the state $|k_2\rangle$ etc. We then simply antisymmetrize and obtain the N-fermion state $|k_1, k_2, \ldots k_N\rangle$. Due to the antisymmetry none of the k_j may coincide.

The next step now is to introduce the occupation number space (N-space) representation for these states. Thus, we define a state $|n_0, n_1, \ldots, n_\infty\rangle$ where each n_k is either 0 or 1. If $n_k = 0$ it means there is no particle in the state $|k\rangle$, whereas if $n_k = 1$ there is exactly one particle in the state $|k\rangle$. The antisymmetry of fermion states restricts the occupation numbers n_k to the two values (0,1). On the other hand the antisymmetry further gives a one to one relation between an N-fermion state $|k_1, k_2, \ldots k_N\rangle$ and a state $|n_0, n_1, \ldots, n_\infty\rangle$ where

$$\sum_{k=0}^{\infty} n_k = N \quad .$$

For example, the 2-particle state $|0, 3\rangle$ corresponding to

$$|0, 3\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle|3\rangle - |3\rangle|0\rangle\right) \tag{20.8.80}$$

where the first state on the right of (20.8.80) refers to particle 1 and the second state to particle 2 is now written as $|1, 0, 0, 1, 0, 0 \ldots\rangle$. This simply says that there is one particle in the ground state $k = 0$ and one particle in the excited state $k = 3$. Notice again that due to the Pauli exclusion principle we can have at most one particle in a given state.

We now assume that the one-particle states $\{|k\rangle\ ,\ k = 0, 1, 2, \ldots\}$ are eigenstates of a one-particle Hamiltonian $H(x_j)$ such that the N-particle Hamiltonian is

$$H = \sum_{j=1}^{N} H(x_j)\ . \tag{20.8.81}$$

Then we have that

$$H(x_j)|k\rangle = E_k|k\rangle \tag{20.8.82}$$

and the state $|n_0, n_1, \ldots, n_\infty\rangle$ is an eigenstate of H

$$H|n_0, n_1, n_1, \ldots\rangle = \left(\sum_{k=0}^{\infty} n_k E_k\right) |n_0, n_1, n_1, \ldots\rangle\ . \tag{20.8.83}$$

We now again introduce number operators N_k that state whether the state $|k\rangle$ is occupied or not. Thus, they must have the eigenvalues 0,1. This requires that

$$N_k^2 = N_k\ . \tag{20.8.84}$$

We next try writing the number operators in terms of annihilation and creation operators

$$N_k = a_k^\dagger a_k \tag{20.8.85}$$

This time, however, a_k and $a_k^\dagger$ cannot satisfy the commutation relations (20.7.70) and (20.7.71) for otherwise we would have boson operators as before and the number operators would have all non-negative integers as eigenvalues and would therefore not satisfy (20.8.84). Furthermore the states produced would be symmetric. To ensure that a state $|n_0, n_1, n_1, \ldots\rangle$ can still be written in the form

$$|n_0, n_1, \ldots\rangle = \frac{(a_0^\dagger)^{n_0}}{\sqrt{n_0!}} \frac{(a_1^\dagger)^{n_1}}{\sqrt{n_1!}} \cdots |0, 0, 0, \ldots\rangle \tag{20.8.86}$$

with not more than one particle in a given state requires that

$$(a_k^\dagger)^2 = 0\ . \tag{20.8.87}$$

This immediately implies that

$$(a_k)^2 = 0\ . \tag{20.8.88}$$

Furthermore the two-particle state

$$|0, \ldots \underset{k}{,1,}\ 0, \ldots, 0,\ \underset{j}{1,}\ 0\ , \ldots\ \rangle = a_k^\dagger a_j^\dagger |0, 0, 0, \ldots\rangle \tag{20.8.89}$$

is antisymmetric in k and j. This requires that

$$a_k^\dagger a_j^\dagger = -a_j^\dagger a_k^\dagger \tag{20.8.90}$$

or

$$a_k^\dagger a_j^\dagger + a_j^\dagger a_k^\dagger = 0\ . \tag{20.8.91}$$

Again the hermitian adjoint of this equation yields

$$a_k a_j + a_j a_k = 0\ . \tag{20.8.92}$$

Equations (20.8.91) and (20.8.92) imply (20.8.87) and (20.8.88). If we now further postulate that

$$a_k^\dagger a_j + a_j a_k^\dagger = \delta_{kj} \tag{20.8.93}$$

it is a simple matter to verify (20.8.84). Thus, we need

$$N_k^2 = N_k\ .$$

But this reads

$$a_k^\dagger a_k a_k^\dagger a_k = a_k^\dagger a_k (a_k^\dagger a_k + a_k a_k^\dagger) = a_k^\dagger a_k \tag{20.8.94}$$

as required. The first equality follows from $a_k^2 = 0$ and the second from (20.8.93).

Thus, fermion creation and annihilation operators satisfy the anti-commutation relations given by (20.8.91) - (20.8.93). We henceforth write these equations as

$$[a_k, a_j]_+ = [a_k^\dagger, a_j^\dagger]_+ = 0 \tag{20.8.95}$$

and

$$[a_k^\dagger, a_j]_+ = \delta_{kj} \tag{20.8.96}$$

where

$$[A, B]_+ = AB + BA \tag{20.8.97}$$

is called the *anti-commutator* of A and B.

Computations for fermions may now be carried out in the same manner as for bosons except that the creation and annihilation operators satisfy anti-commutation instead of commutation relations. Thus, the Hamiltonian (20.8.81) may now be written

$$H = \sum_{k=0}^{\infty} E_k a_k^\dagger a_k\ . \tag{20.8.98}$$

An N-particle state is now picked out by restricting ourselves to that subspace of occupation number space for which

$$\sum_{k=0}^{\infty} n_k = N\ . \tag{20.8.99}$$

Corresponding to this it is convenient to define a total particle number operator N such that

$$\mathsf{N} = \sum N_k = \sum a_k^\dagger a_k \ . \tag{20.8.100}$$

An N-particle state then belongs to the subspace for which the total number operator N has the eigenvalue N.

Again an arbitrary state $|\Psi\rangle$ may be expanded in a basis of occupation number states. Thus,

$$\Psi = \sum |n_0, n_1, \ldots\rangle\langle n_0, n_1, \ldots|\Psi\rangle \tag{20.8.101}$$

just as for the case of bosons. The completeness and orthonormality relations are also the same as for bosons namely (20.7.75) and (20.7.79). The expansion coefficients $\langle n_0, n_1, \ldots|\Psi\rangle$ are just the Slater determinant (20.5.57) and give the probability amplitude for finding n_0 particles in the state $|0\rangle$, n_1 particles in the state $|1\rangle$, etc. Throughout, of course, all the n_k are either 0 or 1.

20.9 Field Operators in the Schrödinger Picture

In this section we treat the Fermi and Bose case at the same time. Where necessary we indicate the differences. We begin by defining the field operators $\psi(\mathbf{x})$, $\psi^\dagger(\mathbf{x})$. Although we use the same symbols as previously used for wavefunctions, $\psi(\mathbf{x})$ and $\psi^\dagger(\mathbf{x})$ now represent operators. This is the conventional notation.

$$\psi(\mathbf{x}) = \sum_{k=0}^{\infty} \langle \mathbf{x}|k\rangle a_k \tag{20.9.102}$$

$$\psi^\dagger(\mathbf{x}) = \sum_{k=0}^{\infty} \langle k|\mathbf{x}\rangle a_k^\dagger \ . \tag{20.9.103}$$

We then have the following (anti)commutation relations

$$\begin{aligned}
[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})]_\pm &= \sum_{j,k=0}^{\infty} \langle \mathbf{x}|j\rangle\langle k|\mathbf{y}\rangle[a_j, a_k^\dagger]_\pm \\
&= \sum_{k=0}^{\infty} \langle \mathbf{x}|k\rangle\langle k|\mathbf{y}\rangle \\
&= \delta(\mathbf{x} - \mathbf{y}) \ .
\end{aligned} \tag{20.9.104}$$

Also

$$[\psi(\mathbf{x}), \psi(\mathbf{y})]_\pm = [\psi^\dagger(\mathbf{x}), \psi^\dagger(\mathbf{y})]_\pm = 0 \ . \tag{20.9.105}$$

Throughout this section, the upper sign applies to fermions and the lower sign to bosons.

The inverse of eqns. (20.9.102) and (20.9.103) is

$$a_k = \int \langle k|\mathbf{x}\rangle \psi(\mathbf{x}) d^3x \tag{20.9.106}$$

$$a_k^\dagger = \int \langle \mathbf{x}|k\rangle \psi^\dagger(\mathbf{x}) d^3x \tag{20.9.107}$$

where we have used the closure or completeness condition

$$\int |\mathbf{x}\rangle\langle\mathbf{x}| d^3x = 1 \,. \tag{20.9.108}$$

The total number operator N may now also be expressed in terms of the field operators. Thus, we have

$$\begin{aligned} \mathsf{N} &= \sum_{k=0}^{\infty} a_k^\dagger a_k \\ &= \sum_{k=0}^{\infty} \int d^3x\, d^3y\, \langle\mathbf{x}|k\rangle\langle k|\mathbf{y}\rangle \psi^\dagger(\mathbf{x})\psi(\mathbf{y}) \\ &= \int d^3x\, \psi^\dagger(\mathbf{x})\psi(\mathbf{x}) \quad . \end{aligned} \tag{20.9.109}$$

The operator $\psi^\dagger(\mathbf{x})\psi(\mathbf{x})$ can therefore be interpreted as a *particle density operator.* To further obtain an interpretation of the field operators we calculate their *commutator* with N for both fermion and boson operators.

$$\begin{aligned} [\mathsf{N}, \psi^\dagger(\mathbf{x})] &= \int d^3y\, \{\psi^\dagger(\mathbf{y})\psi(\mathbf{y})\psi^\dagger(\mathbf{x}) - \psi^\dagger(\mathbf{x})\psi^\dagger(\mathbf{y})\psi(\mathbf{y})\} \\ &= \int d^3y\, \psi^\dagger(\mathbf{y})\delta(\mathbf{x}-\mathbf{y}) \\ &= \psi^\dagger(\mathbf{x}) \quad . \end{aligned} \tag{20.9.110}$$

This result holds for both fermions and bosons and identifies $\psi^\dagger(\mathbf{x})$ as a creation operator for those particles for which N is a number operator. Similarly we get that

$$[\mathsf{N}, \psi(\mathbf{x})] = -\psi(\mathbf{x}) \tag{20.9.111}$$

and that $\psi(\mathbf{x})$ is an annihilation operator.

We now define a *vacuum* or *no-particle* state $|\Omega\rangle$ such that

$$\psi(\mathbf{x})|\Omega\rangle = 0 \,. \tag{20.9.112}$$

This implies that

$$a_k|\Omega\rangle = 0 \tag{20.9.113}$$

for all k and hence this vacuum $|\Omega\rangle$ is the same as the state $|0,0,\ldots\rangle$ of occupation number space. That this is a no-particle state is further demonstrated by the fact that

$$\mathsf{N}|\Omega\rangle = 0\ . \tag{20.9.114}$$

It then follows that $\psi^\dagger(\mathbf{x})|\Omega\rangle$ is a one-particle state, since using (20.9.110) we obtain

$$\mathsf{N}\psi^\dagger(\mathbf{x})|\Omega\rangle = \{\psi^\dagger(\mathbf{x})\mathsf{N} + \psi^\dagger(\mathbf{x})\}\,|\Omega\rangle = 1\cdot\psi^\dagger(\mathbf{x})|\Omega\rangle\ . \tag{20.9.115}$$

The probability amplitude for finding the particle, which is in the state $\psi^\dagger(\mathbf{x})|\Omega\rangle$, in the state $|\alpha\rangle$ is $\langle\alpha|\psi^\dagger(\mathbf{x})|\Omega\rangle$. Thus, if $|\alpha\rangle = |\mathbf{y}\rangle$ we get

$$\begin{aligned}\langle\mathbf{y}|\psi^\dagger(\mathbf{x})|\Omega\rangle &= \langle\Omega|\psi(\mathbf{y})\psi^\dagger(\mathbf{x})|\Omega\rangle\\ &= \delta(\mathbf{x}-\mathbf{y})\langle\Omega|\Omega\rangle\\ &= \delta(\mathbf{x}-\mathbf{y})\end{aligned} \tag{20.9.116}$$

showing that $\psi^\dagger(\mathbf{x})|\Omega\rangle$ is the state for a particle located at the point $\mathbf{x}$. Thus, $\psi^\dagger(\mathbf{x})$ creates a particle localized at the point $\mathbf{x}$. Such states are not normalized and must be turned into wave-packets (they must be smeared with a test function) before they belong to Hilbert space. That is, we find

$$\langle\Omega|\psi(\mathbf{x})|\psi^\dagger(\mathbf{x})|\Omega\rangle = \delta(0)\ . \tag{20.9.117}$$

To form wave-packets we can either smear the states

$$|\mathbf{x}\rangle = |\psi^\dagger(\mathbf{x})|\Omega\rangle \tag{20.9.118}$$

to get the states

$$|f\rangle = \int d^3x\, f(\mathbf{x})|\mathbf{x}\rangle = \int d^3x\, f(\mathbf{x})|\psi^\dagger(\mathbf{x})|\Omega\rangle \tag{20.9.119}$$

or else we smear the operators to form smeared operators

$$\psi^\dagger(f) = \int d^3x\, f(\mathbf{x})\psi^\dagger(\mathbf{x}) \tag{20.9.120}$$

from which we form states

$$|f\rangle = \psi^\dagger(f)|\Omega\rangle = \int d^3x\, f(\mathbf{x})\psi^\dagger(\mathbf{x})|\Omega\rangle\ . \tag{20.9.121}$$

In either case the resultant states have a finite norm if $f(\mathbf{x})$ is square integrable since

$$\begin{aligned}\langle f|f\rangle &= \int d^3x\, d^3y\, f^*(\mathbf{x})f(\mathbf{y})\langle\Omega|\psi(\mathbf{x})\psi^\dagger(\mathbf{y})|\Omega\rangle\\ &= \int d^3x\, |f(\mathbf{x})|^2\ .\end{aligned} \tag{20.9.122}$$

The smeared operators may be viewed as creation and annihilation operators for particles localized according to the wave-packets $f(\mathbf{x})$. They satisfy the following (anti) commutation relations:

$$[\psi(f),\psi^\dagger(g)]_\pm = \int d^3x\, f^*(\mathbf{x})g(\mathbf{x}) \tag{20.9.123}$$

$$[\psi(f),\psi(g)]_\pm = [\psi^\dagger(f),\psi^\dagger(g)]_\pm = 0 \ . \tag{20.9.124}$$

It is, therefore, quite natural to treat the probability amplitude $\langle\alpha|\psi^\dagger(\mathbf{x})|\Omega\rangle$ as a wavefunction $F^*_\alpha(\mathbf{x})$, and the corresponding amplitude $\langle\Omega|\psi(\mathbf{x})|\alpha\rangle$ as the wavefunction $F_\alpha(\mathbf{x})$. The reason for our choice of complex conjugates will become clear a little later.

Similarly a two particle wavefunction in configuration space is then given by $\langle\Omega|\psi(\mathbf{x})\psi(\mathbf{y})|\alpha_1,\alpha_2\rangle$ and an N-particle wavefunction in configuration space is given by $\langle\Omega|\psi(\mathbf{x_1})\psi(\mathbf{x_2})\ldots\psi(\mathbf{x_N})|\alpha_1,\alpha_2,\ldots,\alpha_N\rangle$. We shall use these results very soon to show how the second quantized formalism in terms of field operators may be used to describe N-particle Schrödinger equations.

20.10 Representation of Operators

In practice we know the representation of operators in configuration space. We now first rewrite them in the occupation number space representation. This allows us to express them in terms of the field operators.

Thus, suppose we are given a local operator F in configuration space. We assume that this operator does not change the number of particles. In that case its configuration space representation is

$$\begin{aligned} &\langle\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N|F|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_M\rangle \\ = \ &\delta_{NM}F^N(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N)\prod_{j=1}^{N}\delta(\mathbf{x}_j-\mathbf{y}_j) \ . \end{aligned} \tag{20.10.125}$$

We now use the completeness of the states $|\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N\rangle$ to write

$$\begin{aligned} &\langle n_0,n_1,\ldots|F|m_0,m_1,\ldots\rangle = \int d^3x_1\ldots d^3x_N\, d^3y_1\ldots d^3y_N \\ \times\ &\langle n_0,n_1,\ldots|\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N\rangle\langle\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N| \\ &F|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_N\rangle\langle\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_N|m_0,m_1,\ldots\rangle \\ = \ &\int d^3x_1\ldots d^3x_N\, d^3y_1\ldots d^3y_N \\ \times\ &\langle n_0,n_1,\ldots|\psi^\dagger(\mathbf{x}_N)\ldots\psi^\dagger(\mathbf{x}_1)|\Omega\rangle\langle\Omega|\psi(\mathbf{x}_1)\psi(\mathbf{x}_2)\ldots\psi(\mathbf{x}_N) \\ &F|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_N\rangle\langle\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_N|m_0,m_1,\ldots\rangle \end{aligned} \tag{20.10.126}$$

or

$$\langle n_0,n_1,\ldots|F|m_0,m_1,\ldots\rangle = \frac{1}{N!}\int d^3x_1\ldots d^3x_N\langle n_0,n_1,\ldots|$$

$$|\psi^\dagger(\mathbf{x}_N)\ldots\psi^\dagger(\mathbf{x}_1)|\Omega\rangle F^N(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N)$$
$$\langle\Omega|\psi(\mathbf{x}_1)\psi(\mathbf{x}_2)\ldots\psi(\mathbf{x}_N)|m_0,m_1,\ldots\rangle \tag{20.10.127}$$

where we used that

$$\langle\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{\mathbf{N}N}|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_N\rangle = \prod_{j=1}^{N}\delta(\mathbf{x}_j-\mathbf{y}_j) \tag{20.10.128}$$

and also integrated out all the $\mathbf{y}_j$ variables. Furthermore,

$$\sum n_i = \sum m_i = N \ . \tag{20.10.129}$$

Now $|\psi(\mathbf{x}_1)\psi(\mathbf{x}_2)\ldots\psi(\mathbf{x}_N)|m_0,m_1,\ldots\rangle$ with $\sum m_i = N$ has a non-vanishing inner product only with the no-particle state. We may therefore replace the intermediate state $|\Omega\rangle\langle\Omega|$ with a sum over all possible intermediate states, or

$$1 = \sum_{\{n\}}|n_0,n_1,\ldots\rangle\langle n_0,n_1,\ldots| \tag{20.10.130}$$

to get

$$\langle n_0,n_1,\ldots|F|m_0,m_1,\ldots\rangle = \frac{1}{N!}\int d^3x_1\ldots d^3x_N\langle n_0,n_1,\ldots|$$
$$\psi^\dagger(\mathbf{x}_N)\ldots\psi^\dagger(\mathbf{x}_1)F^N(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N)\psi(\mathbf{x}_1)\ldots\psi(\mathbf{x}_N)$$
$$|m_0,m_1,\ldots\rangle \ . \tag{20.10.131}$$

To proceed further we first consider the Hamiltonian

$$H = \sum_{j=1}^{N}H(\mathbf{x}_j) \tag{20.10.132}$$

consisting of a sum of one-particle Hamiltonians. In this case a considerable simplification occurs. This can be seen by taking the terms in the above sum, one at a time. The last term yields

$$\begin{aligned}
&\frac{1}{N!}\int d^3x_1\ldots d^3x_N\langle n_0,n_1,\ldots|\psi^\dagger(\mathbf{x}_N)\ldots\psi^\dagger(\mathbf{x}_1)H(\mathbf{x}_N)\\
&\psi(\mathbf{x}_1)\ldots\psi(\mathbf{x}_N)|m_0,m_1,\ldots\rangle\\
=\ &\frac{1}{N!}\int d^3x_1\ldots d^3x_N\,\langle n_0,n_1,\ldots|\psi^\dagger(\mathbf{x}_N)H(\mathbf{x}_N)\\
&\psi^\dagger(\mathbf{x}_{N-1})\ldots\psi^\dagger(\mathbf{x}_1)\psi(\mathbf{x}_1)\ldots\psi(\mathbf{x}_N)|m_0,m_1,\ldots\rangle
\end{aligned} \tag{20.10.133}$$

where as before $\sum n_i = \sum m_i = N$. The integral over d^3x_1 produces the number operator $\int d^3x_1\psi^\dagger(\mathbf{x}_1)\psi(\mathbf{x}_1)$ applied to the one-particle state

$$\psi(\mathbf{x}_2)\ldots\psi(\mathbf{x}_N)|m_0,m_1,\ldots\rangle$$

and thus yields 1 times the same state. After that, the integral over d^3x_2 again produces the number operator, namely $\int d^3x_2\psi^\dagger(\mathbf{x}_2)\psi(\mathbf{x}_2)$ applied to the two-particle state

$$\psi(\mathbf{x}_3)\ldots\psi(\mathbf{x}_N)|m_0,m_1,\ldots\rangle$$

and yields 2 times this state. Proceeding in this fashion we eventually get

$$\begin{aligned}&\frac{1}{N!}\int d^3x_1\ldots d^3x_N\langle n_0,n_1,\ldots|\psi^\dagger(\mathbf{x}_N)\ldots\psi^\dagger(\mathbf{x}_1)H(\mathbf{x}_N)\psi(\mathbf{x}_1)\ldots\psi(\mathbf{x}_N)\\ &|m_0,m_1,\ldots\rangle\\ &= \frac{1}{N}\int d^3x_N\,\langle n_0,n_1,\ldots|\psi^\dagger(\mathbf{x}_N)H(\mathbf{x}_N)\psi(\mathbf{x}_N)|m_0,m_1,\ldots\rangle\ . \qquad (20.10.134)\end{aligned}$$

Now every term in the sum can be brought to this form by commuting, or anticommuting all the $\psi^\dagger(\mathbf{x}_j)H(\mathbf{x}_j)$ to the extreme left and every $\psi(\mathbf{x}_j)$ to the extreme right. Hence we get that

$$\begin{aligned}&\langle n_0,n_1,\ldots|H|m_0,m_1,\ldots\rangle\\ &= \frac{1}{N!}\sum_{j=1}^{N}\int d^3x_j\langle n_0,n_1,\ldots|\psi^\dagger(\mathbf{x}_j)H(\mathbf{x}_j)\psi(\mathbf{x}_j)|m_0,m_1,\ldots\rangle\\ &= \int d^3x\,\langle n_0,n_1,\ldots|\psi^\dagger(\mathbf{x})H(\mathbf{x})\psi(\mathbf{x})|m_0,m_1,\ldots\rangle\ . \qquad (20.10.135)\end{aligned}$$

Thus, the occupation number space representation of the Hamiltonian corresponding to the sum of one-particle Hamiltonians (20.2.4) is given by

$$\begin{aligned}\mathsf{H} &= \int d^3x\,\psi^\dagger(\mathbf{x})H(\mathbf{x})\psi(\mathbf{x})\\ &= \int d^3x\,\psi^\dagger(\mathbf{x})\left[-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{x})\right]\psi(\mathbf{x})\ . \qquad (20.10.136)\end{aligned}$$

To complete our treatment we also need to consider a general N-body interaction. Fortunately, it appears that for systems whose interactions may be described by potentials, the most general interaction involves a sum of two-body interactions of the form

$$V(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N)=\frac{1}{2}\sum_{j\neq k}^{N}V(\mathbf{x}_j,\mathbf{x}_k)\ . \qquad (20.10.137)$$

No intrinsic three-body forces or higher have so far been found necessary for a description of actual systems.

Applying the results of equation (20.10.131) to the interaction (20.10.137) we obtain in occupation number space the following operator (see problem 20.8)

$$\mathsf{V}=\frac{1}{2}\int d^3x\,d^3y\,\psi^\dagger(\mathbf{x})\psi^\dagger(\mathbf{y})V(\mathbf{x},\mathbf{y})\psi(\mathbf{y})\psi(\mathbf{x})\ . \qquad (20.10.138)$$

The order of the last two operators is deliberately reversed to remove the self-interaction $V(\mathbf{x},\mathbf{x})$ which does not occur in (20.10.137).

Thus, a Hamiltonian containing an external potential $V_0(\mathbf{x})$ as well as two-body interactions $V(\mathbf{x},\mathbf{y})$ would have the second-quantized form

$$\begin{aligned} \mathsf{H} &= \int d^3x\,\psi^\dagger(\mathbf{x})\left[-\frac{\hbar^2}{2m}\nabla^2 + V_0(\mathbf{x})\right]\psi(\mathbf{x}) \\ &+ \frac{1}{2}\int d^3x\,d^3y\,\psi^\dagger(\mathbf{x})\psi^\dagger(\mathbf{y})V(\mathbf{x},\mathbf{y})\psi(\mathbf{y})\psi(\mathbf{x})\ . \end{aligned} \tag{20.10.139}$$

20.11 Heisenberg Picture

We next transform to the Heisenberg picture and obtain the equations of motion for the field operators. Many theoretical discussions commence with these field equations, which are known as the Heisenberg equations. They are, as we shall see, a sophisticated summary of the Planck frequency condition $E = h\nu$.

The field operators we have used so far have all been in the Schrödinger picture. To distinguish them from the Heisenberg picture operators, we are about to introduce, we label them with a subscript s. Thus, the states in the Heisenberg picture are related to the states in the Schrödinger picture through

$$|\Psi\rangle = \exp(i\mathsf{H}t/\hbar)|\Psi_s(t)\rangle \tag{20.11.140}$$

and the operators are related by

$$\psi(\mathbf{x},t) = \exp(i\mathsf{H}t/\hbar)\psi_s(\mathbf{x})\exp(-i\mathsf{H}t/\hbar)\ . \tag{20.11.141}$$

Since these transformations are unitary they preserve the equal time (anti) commutation relations. Thus,

$$[\psi(\mathbf{x},t),\psi^\dagger(\mathbf{y},t)]_\pm = \delta(\mathbf{x}-\mathbf{y}) \tag{20.11.142}$$

and

$$[\psi(\mathbf{x},t),\psi(\mathbf{y},t)]_\pm = [\psi^\dagger(\mathbf{x},t),\psi^\dagger(\mathbf{y},t)]_\pm = 0\ . \tag{20.11.143}$$

By a straightforward differentiation of equation (20.11.141) we obtain the *Heisenberg equation of motion* for the field operators

$$\begin{aligned} i\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t} &= [\psi(\mathbf{x},t),\mathsf{H}] \\ i\hbar\frac{\partial\psi^\dagger(\mathbf{x},t)}{\partial t} &= [\psi^\dagger(\mathbf{x},t),\mathsf{H}]\ . \end{aligned} \tag{20.11.144}$$

If we consider the Hamiltonian (20.10.139) then since H is time independent we get

$$\begin{aligned} &[\psi(\mathbf{x},t),\mathsf{H}] = \int d^3y\,[\psi(\mathbf{x},t),\psi^\dagger(\mathbf{y},t)]\left\{-\frac{\hbar^2}{2m}\nabla^2 + V_0(\mathbf{y})\right\} + \\ &\frac{1}{2}\int d^3y\,d^3z\,[\psi(\mathbf{x},t),\psi^\dagger(\mathbf{y},t)\psi^\dagger(\mathbf{z},t)\psi(\mathbf{z},t)\psi(\mathbf{y},t)]\,V(\mathbf{y},\mathbf{z}). \end{aligned} \tag{20.11.145}$$

This yields the following field equation

$$\begin{aligned} i\hbar\frac{\partial\psi}{\partial t} &= \left\{-\frac{\hbar^2}{2m}\nabla^2 + V_0(\mathbf{x})\right\}\psi(\mathbf{x},t) \\ &+ \int d^3y\,\psi^\dagger(\mathbf{y},t)V(\mathbf{y},\mathbf{x})\psi(\mathbf{y},t)\psi(\mathbf{x},t)\;. \end{aligned} \tag{20.11.146}$$

To illustrate how this is obtained we consider the last term in equation (20.11.145). As before we treat fermions and bosons simultaneously; the upper sign always applies to fermions. We first note that

$$[\psi(\mathbf{x},t),\psi(\mathbf{z},t)\psi(\mathbf{y},t)] = 0 \tag{20.11.147}$$

both for fermions and bosons. Therefore the commutator in the last term in (20.11.145) reduces to

$$[\psi(\mathbf{x},t),\psi^\dagger(\mathbf{y},t)\psi^\dagger(\mathbf{z},t)]\psi(\mathbf{z},t)\psi(\mathbf{y},t)\;. \tag{20.11.148}$$

But,

$$\begin{aligned} &[\psi(\mathbf{x},t),\psi^\dagger(\mathbf{y},t)\psi^\dagger(\mathbf{z},t)] = \psi(\mathbf{x},t)\psi^\dagger(\mathbf{y},t)\psi^\dagger(\mathbf{z},t) - \psi^\dagger(\mathbf{z},t)\psi^\dagger(\mathbf{y},t)\psi(\mathbf{x},t) \\ = &\;\left\{\psi(\mathbf{x},t)\psi^\dagger(\mathbf{y},t) \pm \psi^\dagger(\mathbf{y},t)\psi(\mathbf{x},t)\right\}\psi^\dagger(\mathbf{z},t) \\ \mp &\;\psi^\dagger(\mathbf{y},t)\left\{\psi(\mathbf{x},t)\psi^\dagger(\mathbf{z},t) \pm \psi^\dagger(\mathbf{z},t)\psi(\mathbf{x},t)\right\} \\ = &\;\delta(\mathbf{x}-\mathbf{y})\psi^\dagger(\mathbf{z},t) \mp \delta(\mathbf{x}-\mathbf{z})\psi^\dagger(\mathbf{y},t)\;. \end{aligned} \tag{20.11.149}$$

Combining all these results and integrating out the delta functions yields

$$\begin{aligned} &\frac{1}{2}\int d^3y\,d^3z\,[\psi(\mathbf{x},t),\psi^\dagger(\mathbf{y},t)\psi^\dagger(\mathbf{z},t)\psi(\mathbf{z},t)\psi(\mathbf{y},t)]V(\mathbf{y},\mathbf{z}) \\ = &\;\frac{1}{2}\int d^3y\,\psi^\dagger(\mathbf{y},t)V(\mathbf{y},\mathbf{x})\left\{\psi(\mathbf{y},t)\psi(\mathbf{x},t) \mp \psi(\mathbf{x},t)\psi(\mathbf{y},t)\right\} \\ = &\;\int d^3y\,\psi^\dagger(\mathbf{y},t)V(\mathbf{y},\mathbf{x})\psi(\mathbf{y},t)\psi(\mathbf{x},t)\;. \end{aligned} \tag{20.11.150}$$

To obtain the first equality we used Newton's third law in the form $V(\mathbf{x},\mathbf{y}) = V(\mathbf{y},\mathbf{x})$.

Equation (20.11.146) occurs whenever we have two-body interactions. Thus, in a sense it contains almost all of solid state physics. The problem is that it is not only an equation for operators and thus represents an infinite number of scalar equations, but it is also non-linear. There is therefore no hope of an analytic solution of this equation except in the most trivial cases. In practice it is frequently more convenient to work with the annihilation and creation operators. In that case the Hamiltonian (20.10.139) becomes

$$\mathsf{H} = \sum_{m,n} E_{m,n}a_m^\dagger a_n + \sum_{k,l,m,n} V_{k,l,m,n}a_k^\dagger a_l^\dagger a_m a_n \tag{20.11.151}$$

where

$$E_{m,n} = \int d^3x\,\langle m|\mathbf{x}\rangle\left[-\frac{\hbar^2}{2m}\nabla^2 + V_0(\mathbf{x})\right]\langle\mathbf{x}|n\rangle \tag{20.11.152}$$

and

$$V_{k,l,m,n} = \int d^3x\, d^3y\, \langle k|\mathbf{x}\rangle\langle l|\mathbf{x}\rangle V(\mathbf{x},\mathbf{y})\langle \mathbf{x}|m\rangle\langle \mathbf{y}|n\rangle \ . \tag{20.11.153}$$

The states $|n\rangle$ here denote one-particle basis states. If they are chosen as eigenstates of the single particle Hamiltonian

$$H_0 = -\frac{\hbar^2}{2m}\nabla^2 + V_0(\mathbf{x}) \tag{20.11.154}$$

with eigenvalues E_n, then equation (20.11.152) reduces to

$$E_{m,n} = E_n \delta_{n,m} \ . \tag{20.11.155}$$

In that case

$$\mathsf{H} = \sum_{m=0} E_m a_m^\dagger a_m + \sum_{k,l,m,n} V_{k,l,m,n} a_k^\dagger a_l^\dagger a_m a_n \ . \tag{20.11.156}$$

As a final consideration we assume that we have eigenstates of energy E of the Hamiltonian (20.11.151) of our system. Denoting these states by $|E\rangle$ and taking matrix elements in this basis of the Heisenberg equation of motion (20.11.144) for the field operator we get

$$i\hbar\langle E|\frac{\partial\psi(\mathbf{x},t)}{\partial t}|E'\rangle = (E' - E)\langle E|\psi(\mathbf{x},t)|E'\rangle \ . \tag{20.11.157}$$

If we furthermore Fourier decompose $\psi(\mathbf{x},t)$ as

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \psi(\mathbf{x},\omega)\, e^{-i\omega t}\, d\omega \tag{20.11.158}$$

we find that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} [\hbar\omega - (E' - E)]\langle E|\psi(\mathbf{x},\omega)|E'\rangle e^{-i\omega t}\, d\omega = 0 \tag{20.11.159}$$

and we have that

$$\langle E|\psi(\mathbf{x},\omega)|E'\rangle = 0 \quad \text{unless} \quad E' - E = \hbar\omega \ . \tag{20.11.160}$$

Thus, we have rederived the Planck frequency relation from the Heisenberg equation of motion.

20.12 Problems

20.1 Using the Pauli exclusion principle, determine the maximum number of electrons in any energy level n of an atom. Neglect the interactions between the electrons.

20.2 Two electrons are in the d-state ($l = 2$) of an helium atom. Write all permitted spin wavefunctions for this case and state what all the possible total angular momentum values are.

20.3 Two identical spin 1/2 particles collide.
a) Assume that the wavefunction for the two particles after the collision corresponds to a relative orbital angular momentum with $l = 0$. Write out the possible spinor wavefunctions.
b) Repeat part a) if $l = 1$.

20.4 Consider the collision of two identical spin 0 particles. Separate the wavefunction into centre of mass and relative coordinates. Discuss the symmetry required of the relative wave function. Use this to obtain the modifications required for the scattering amplitude and hence an expression for the differential cross-section.

20.5 Repeat problem 20.4 for two identical spin 1/2 particles. Again, separate the wavefunction into centre of mass and relative coordinates as well as spin coordinates. Assume the particles scatter from a spin-independent potential.
a) Discuss the symmetry required of the relative wavefunction if the scattering occurs in a singlet ($s = 0$) state. Use this to obtain the modifications required for the scattering amplitude and hence an expression for the differential cross-section.
b) Repeat part a) if the scattering occurs in a triplet $s = 1$ state.

20.6 Consider a system of non-interacting bosons and write the Hamiltonian in the form

$$\mathsf{H} = \sum_{k=0}^{\infty} \hbar\omega a_k^\dagger a_k \ .$$

Find an explicit expression for

$$\exp[i\mathsf{H}t/\hbar]\, \psi(\mathbf{x}, t)\, \exp[-i\mathsf{H}t/\hbar]$$

where

$$\psi_s(\mathbf{x}) = \sum_{k=0}^{\infty} \langle \mathbf{x}|k\rangle a_k \ .$$

Hint: Expand the second exponential and commute a_k through, showing that

$$a_k\, e^{\lambda a_k^\dagger a_k} = e^{\lambda(a_k^\dagger a_k + 1)}\, a_k \ .$$

20.7 Repeat problem 21.6 for fermions.

20.8 Obtain equation (20.10.138) for the occupation number space representation of the interaction specified by equation (20.10.137).

20.9 Show that the Hamiltonian

$$H = \sum_k \left[E_k a_k^\dagger a_k + \lambda_k a_k^\dagger a_k^\dagger + \lambda_k^* a_k a_k \right] \quad , \quad E_k > 2|\lambda_k|$$

can be diagonalized if a_k, $a_k^\dagger$ are bose operators.
Hint: introduce operators

$$b_k = u_k a_k^\dagger + v_k a_k$$

and choose the constants u_k and v_k appropriately. What happens if a_k, $a_k^\dagger$ are fermi operators? These are simple examples of Bogoliubov transformations.

20.10 Let A and B be arbitrary operators. Derive the following formula

$$e^{-A}\, B\, e^{A} = \sum_{n=0}^{\infty} \frac{1}{n!} [B, A]_n$$

where,

$$[B, A]_0 = B \quad , \quad [B, A]_{n+1} = [[B, A]_n, A] \, .$$

20.11 Let a_k, $a_k^\dagger$ be fermi operators $k = 1, 2$ and define

$$b_1^\dagger = u a_1^\dagger - v a_2 \quad , \quad b_1 = u^* a_1 - v^* a_2^\dagger$$

$$b_2 = v a_1^\dagger + u a_2 \quad , \quad b_2^\dagger = v^* a_1 + u^* a_2^\dagger$$

with

$$|u|^2 + |v|^2 = 1 \quad , \quad uv^* - u^* v = 0 \, .$$

a) Verify that b_k, $b_k^\dagger$ are fermi operators.
b) Show that for an appropriate choice of the constant c

$$b_1^\dagger = U a_1^\dagger U^\dagger \quad , \quad b_1 = U a_1 U^\dagger$$

$$b_2^\dagger = U a_2^\dagger U^\dagger \quad , \quad b_2 = U a_2 U^\dagger$$

where

$$U = \exp\left\{ c\, a_1^\dagger a_2^\dagger - c^*\, a_2 a_1 \right\}$$

is a unitary operator. This is known as a Bogoliubov transformation.
Hint: Use the result of problem 20.10 .

20.12 a) Diagonalize the Hamiltonian

$$\mathsf{H} = \sum_{k=0}^{\infty} \hbar\omega a_k^\dagger a_k + \frac{1}{2}\sum_k V(k)[a_k a_k + a_k^\dagger a_k^\dagger] \, .$$

where $\omega(k)$ and $V(k)$ are given functions of k and

$$[a_k, a_{k'}^\dagger] = \delta_{k,k'} \quad , \quad [a_k, a_k] = 0 \, .$$

Hint: Use the Bogoliubov transformation

$$b_k = u_k a_k - v_k a_k^\dagger \quad , \quad b_k^\dagger = u_k^* a_k^\dagger - v_k^* a_k$$

with

$$|u_k|^2 - |v_k|^2 = 1 \, .$$

b) Is this Bogoliubov transformation unitary for all $\omega(k)$ and $V(k)$? If not what conditions on $\omega(k)$ and $V(k)$ will guarantee that it is unitary?

20.13 Consider a finite set N of bose operators a_k such that

$$[a_k, a_p^\dagger] = \delta_{k,p} \quad , \quad [a_k, a_{-p}] = 0 \quad , \quad [a_k^\dagger, a_p^\dagger] = 0 \ .$$

Define

$$b_k = \cosh\lambda \, a_k + \sinh\lambda \, a_k^\dagger$$

$$b_k^\dagger = \cosh\lambda \, a_k^\dagger + \sinh\lambda \, a_k \ .$$

a) Find the commutation rules for the b_k, $b_k^\dagger$.
b) Find a unitary operator

$$V_N = \exp(iT_N)$$

in terms of the a_k, $a_k^\dagger$ such that

$$V_N \, a_k \, V_N^\dagger = b_k \ .$$

c) Show that

$$\lim_{N\to\infty} \langle\Psi|V_N|\Phi\rangle = 0$$

where $|\Psi\rangle$, $|\Phi\rangle$ are any states of the form

$$\prod_k a_k^\dagger |0\rangle \ .$$

What does this last result mean?

Bibliography

[20.1] Electrostatics is treated in chapters 1-3 of
J.D. Jackson *Classical Electrodynamics*, John Wiley and Sons, Inc., New York (1962).

[20.2] Y. Sugiura, Z. Physik **45**, 484 (1927).

[20.3] P.M. Morse and H. Feshbach, *Methods of Theoretical Physics*, Vol.1, McGraw-Hill Book Co., Inc., New York (1953).

[20.4] The reader interested in further applications of quantum mechanics to chemistry may consult
W. Kauzmann, *Quantum Chemistry*, Academic Press Inc., New York (1957).

[20.5] There are many ways to approach second quantization. For an approach different from ours see
A.L. Fetter and J.D. Walecka, *Quantum Theory of Many-Particle Systems*, McGraw-Hill Book Co., New York (1971).

Chapter 21

Quantum Statistical Mechanics

21.1 Introduction

In the previous chapter we developed techniques for dealing with systems of particles. If the system consists of a very large number of, say 10^{23}, particles then it is neither possible nor desirable to have exact knowledge of the state of the system. In this case, statistical techniques are required to handle this incomplete knowledge. The procedure here is quite analogous to classical statistical mechanics as developed by Boltzmann and Gibbs.

In quantum statistical mechanics, the probability or statistical concepts enter at two levels. There is the statistical distribution of results of a measurement of an observable on identically prepared systems. This has been the subject of our discussion up until now. These quantum mechanical probabilities add coherently and are described by probability amplitudes. There is also the statistical distribution due to an incomplete knowledge of the state of the system. These are the incoherent probabilities used in classical statistical mechanics. Thus, in equilibrium, this second level of probabilities is determined by an ensemble, as developed by Gibbs. We now develop techniques to incorporate both effects.

The approach we take is more heuristic than rigorous, but it brings out the relevant physical input. We start by dividing the universe into two parts

1) The system of interest to us; this we simply call our "system".

2) The external world.

Thus, for example, we could consider a gas in a container as our system and the rest of the universe including the physical container as the external world. We furthermore assume that our system interacts "weakly" with the external world through the boundaries of the system (walls of the container).

Let $|\psi\rangle$ be a ket describing both our system and the external world,and let $\{|k\rangle\}$ be a complete orthonormal set of kets for our system. In that case we can write

$$|\psi\rangle = \sum_k |c_k\rangle|k\rangle \tag{21.1.1}$$

where the $\{|c_k\rangle\}$ are a complete set of kets for the external world.

Now suppose A is an operator corresponding to an observable of our system. Thus, A operates on the space of kets $\{|k\rangle\}$. An instantaneous expectation value of this observable is given by

$$\frac{\langle\psi|A|\psi\rangle}{\langle\psi|\psi\rangle} .$$

This quantity represents an average result for a large number of identical measurements *performed at the same time.* We now rewrite this expression

$$\frac{\langle\psi|A|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\sum_{n,m}\langle c_n|c_m\rangle\langle n|A|m\rangle}{\sum_n\langle c_n|c_n\rangle} . \tag{21.1.2}$$

We have used here the fact that A corresponds to an observable of our system and does not operate on the states of the external world. Now $\langle\psi|\psi\rangle$ is time independent and therefore so is also $\sum_n\langle c_n|c_n\rangle$.

In a laboratory, we do not perform instantaneous measurements on a many-particle system, but rather time averaged measurements. Thus, we measure

$$\langle\langle A\rangle\rangle = \frac{\overline{\langle\psi|A|\psi\rangle}}{\overline{\langle\psi|\psi\rangle}} = \frac{\sum_{n,m}\overline{\langle c_n|c_m\rangle}\langle n|A|m\rangle}{\sum_n\overline{\langle c_n|c_n\rangle}} . \tag{21.1.3}$$

Here the bar represents a time average. The assumptions of quantum statistical mechanics, when referring to a macroscopic observable of a macroscopic system in thermodynamic equilibrium, are assumptions on the time-averaged coefficients $\overline{\langle c_n|c_m\rangle}$. These coefficients form an object called the *density matrix.*

21.2 The Density Matrix

We begin with some examples to illustrate the previous ideas about the density matrix. These examples conform as closely as possible to classical statistical mechanics.

21.2.1 The Microcanonical Ensemble

In this case we assume that our system consists of N particles in a volume V. We further assume that their energy is fixed between E and $E+\Delta$ where $\Delta << E$. If H is the (N-particle) Hamiltonian of our system we choose as our basis set $\{|k\rangle\}$ the set of eigenkets of H.

$$H|k\rangle = E_k|k\rangle . \tag{21.2.4}$$

Here $|k\rangle$ represents a state for N particles with total energy E_k enclosed in a volume V. We now make the following statistical assumptions.

M1. $\overline{\langle c_n|c_m\rangle} = \begin{cases} 1 & E < E_n < E + \Delta \\ 0 & \text{otherwise} \end{cases}$
Assumption of equal a priori probabilities.

M2. $\overline{\langle c_n|c_m\rangle} = 0$ for $n \neq m$
Assumption of random phases.

We can thus write

$$|\psi\rangle = \sum_n b_n|n\rangle \tag{21.2.5}$$

where the b_n are numbers with random phases such that

$$|b_n|^2 = \begin{cases} 1 & E < E_n < E + \Delta \\ 0 & \text{otherwise} \end{cases}$$

This is a specification by the outside world or heat bath.

We thus get

$$\langle\langle A\rangle\rangle = \frac{\sum_n |b_n|^2\langle n|A|n\rangle}{\sum_n |b_n|^2} \ . \tag{21.2.6}$$

The postulate of random phases implies that the equilibrium state is an incoherent (classical) superposition of eigenstates. The quantities,

$$\frac{|b_n|^2}{\sum_n |b_n|^2}$$

may clearly be considered as classical probabilities p_n. Since this is the case, we can rewrite the above expression so that no reference to phases need occur. Thus,

$$\begin{aligned}
\langle\langle A\rangle\rangle &= \sum_n p_n\langle n|A|n\rangle \\
&= \sum_{n,m}\langle n|A|m\rangle p_n\langle m|n\rangle \\
&= \sum_{n,m}\langle m|n\rangle p_n\langle n|A|m\rangle \\
&= \operatorname{Tr}\left(\sum_n |n\rangle p_n\langle n|A\right) \\
&= \operatorname{Tr}\left(\sum_n A|n\rangle p_n\langle n|\right) \qquad (21.2.7)
\end{aligned}$$

where Tr means "trace" or "sum of the diagonal elements" of the operator following. This expression can also be written as

$$\langle\langle A\rangle\rangle = \mathrm{Tr}(A\rho) \tag{21.2.8}$$

where

$$\rho = \sum_n |n\rangle p_n \langle n| \ . \tag{21.2.9}$$

The operator ρ is called the density matrix. Thus, in our example of the microcanonical ensemble the density matrix is, up to normalization, given by

$$\rho = \sum_{E<E_n<E+\Delta} |n\rangle\langle n| \ . \tag{21.2.10}$$

It is now an easy matter to display the density matrix for other ensembles.

21.2.2 The Canonical Ensemble

In this case we have N particles in a volume V. The total energy however is no longer fixed, instead states are weighted with the Boltzmann factor $e^{-\beta E_n}$ where $\beta = 1/(k_B T)$, with k_B Boltzmann's constant. Thus, in analogy to the microcanonical ensemble we make two statistical assumptions.

C1. $\overline{\langle c_n|c_n\rangle} = e^{-\beta E_n}$.
Assumption of weighting according to the Boltzmann factor

C2. $\overline{\langle c_n|c_m\rangle} = 0$ for $n \neq m$.
Assumption of random phases.

We then obtain for the *unnormalized* density matrix

$$\begin{aligned}\rho &= \sum_n |n\rangle e^{-\beta E_n}\langle n| \\ &= e^{-\beta H}\sum_n |n\rangle\langle n|\end{aligned} \tag{21.2.11}$$

or

$$\rho = \exp(-\beta \mathsf{H}) \ . \tag{21.2.12}$$

One also defines the normalization or *partition function* for N particles Z_N by:

$$Z_N = \mathrm{Tr}\exp(-\beta\mathsf{H}) \ . \tag{21.2.13}$$

This quantity yields the interesting macroscopic or thermodynamic observables. We defer a discussion of this until later.

21.2.3 The Grand Canonical Ensemble

If we wish to consider a system in which even the particle number N is not specified, but only the average number, we can introduce a further "Boltzmann factor" $\exp(\beta\mu\mathsf{N})$ for the particle number. In this case μ is the *chemical potential.* Thus, in this case, we have for the unnormalized density matrix

$$\rho = \exp[-\beta(\mathsf{H} - \mu\mathsf{N})] \ . \tag{21.2.14}$$

This yields a *grand partition function* (normalization of the density matrix)

$$Z_G = \mathrm{Tr}\exp[-\beta(\mathsf{H} - \mu\mathsf{N})] \ . \tag{21.2.15}$$

The connection between statistical mechanics and thermodynamics is made through the thermodynamic functions. This is the same as in classical statistical mechanics. Thus, the internal energy U of a system in the Grand Canonical Ensemble is given by the average value of the Hamiltonian. This means

$$U = Z_G^{-1}\mathrm{Tr}(\rho\mathsf{H}) \ . \tag{21.2.16}$$

Since ρ is given by (21.2.14) it is easily seen that

$$U = -\frac{\partial}{\partial\beta}\ln Z_G(\mu, V, T) + \mu\langle\langle\mathsf{N}\rangle\rangle \ . \tag{21.2.17}$$

This, however, is not the most convenient form for the expression of the internal energy. A better form is obtained by first expressing Z_G as a function of the *fugacity,* volume and temperature. This is done below.

The average number of particles in our system is given by

$$\langle\langle\mathsf{N}\rangle\rangle = Z_G^{-1}\mathrm{Tr}(\rho\mathsf{N}) \ . \tag{21.2.18}$$

To evaluate this expression it is convenient to introduce the fugacity

$$z = e^{\beta\mu} \ . \tag{21.2.19}$$

Then,

$$\rho = \exp(\mu\mathsf{N})\exp-(\beta\mathsf{H}) = z^N\exp-(\beta\mathsf{H}) \ . \tag{21.2.20}$$

Hence we get that the grand partition function is

$$Z_G = \sum_{n=0}^{\infty} z^n Z_n \ . \tag{21.2.21}$$

Combining this result with (21.2.18) we obtain

$$\langle\langle\mathsf{N}\rangle\rangle = z\frac{\partial}{\partial z}\ln Z_G \ . \tag{21.2.22}$$

Equation (21.2.22) is usually inverted to express the chemical potential in terms of the average density of particles. This result is then substituted into equation (21.2.17) to express the internal energy as a function of temperature and density. In fact, by performing Legendre transformations (see section 2.3) it is now possible to compute all other thermodynamic functions. Thus once Z_G is obtained, all thermodynamic quantities are determined in principle.

We simply list the relevant relations. Their derivation is left as an exercise.

$$Z_G = \mathrm{Tr}\left[\exp -\beta(\mathsf{H} - \mu\mathsf{N})\right] = \sum_{n=0}^{\infty} z^n Z_n \ . \tag{21.2.23}$$

The average number $N = \langle\langle \mathsf{N} \rangle\rangle$ of particles in our system (of volume V) is given by

$$N = \langle\langle \mathsf{N} \rangle\rangle = z\frac{\partial}{\partial z} \ln Z_G \ . \tag{21.2.24}$$

The internal energy can also be written as

$$U = -\frac{\partial}{\partial \beta} \ln Z_G(\mu, V, T) + \mu\langle\langle \mathsf{N} \rangle\rangle \ . \tag{21.2.25}$$

Just as in classical statistical mechanics we also have

$$\frac{PV}{k_B T} = \ln Z_G \ . \tag{21.2.26}$$

The three equations (21.2.24), (21.2.25), and (21.2.26) suffice to determine all thermodynamic quantities. Thus, for example, to compute the entropy S one uses (21.2.24) to eliminate z from (21.2.25) and to express the internal energy U as a function of N, V and T. The specific heat at constant volume C_V is now obtained as

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V \ . \tag{21.2.27}$$

This allows us to express the entropy as

$$dS = C_V \frac{dT}{T} \ . \tag{21.2.28}$$

For more details on thermodynamic relations we direct the reader to the references at the end of this chapter.

Next we consider, by way of illustration the simplest possible systems, the ideal gases.

21.3 The Ideal Gases

We now combine the discussion of the previous section with the second quantization techniques developed in sections 7. and 8. of Chapter 20. This allows us to illustrate the usefulness of these techniques as well as to derive the Bose-Einstein and Fermi-Dirac distributions.

The reason, the second quantized formalism is so useful, is that for the grand canonical ensemble the number of particles is not fixed. This makes the occupation number space representation ideally suited for computations in the grand canonical ensemble.

The system we consider consists of an indefinite number of non-interacting (free) particles confined to a fixed volume V. This system can therefore be described by the Hamiltonian

$$\mathsf{H} = \sum_k E_k a_k^\dagger a_k = \sum_k E_k N_k \tag{21.3.29}$$

where as discussed in section 7. and 8. of Chapter 20, the E_k are one-particle energies and we are working in the occupation number space representation. To evaluate

$$Z_G = \mathrm{Tr}\left\{\exp\left[-\beta(\mathsf{H} - \mu\mathsf{N})\right]\right\} \tag{21.3.30}$$

we simply take matrix elements for a complete set of states and sum over the diagonal elements.

$$\begin{aligned} Z_G &= \sum_{n_1 \ldots n_\infty} \langle n_1 \ldots n_\infty | \exp\left[-\beta(\mathsf{H} - \mu\mathsf{N})\right] | n_1 \ldots n_\infty \rangle \\ &= \sum_{n_1 \ldots n_\infty} \langle n_1 \ldots n_\infty | e^{-\beta(\sum_k E_k N_k - \mu \sum_k N_k)} | n_1 \ldots n_\infty \rangle \ . \end{aligned} \tag{21.3.31}$$

We have simply used (21.3.29) and the definition

$$\mathsf{N} = \sum_k N_k \tag{21.3.32}$$

of the total number operator.

Using,

$$N_k | n_1 \ldots n_\infty \rangle = n_k | n_1 \ldots n_\infty \rangle \tag{21.3.33}$$

the expression for the grand partition function becomes

$$Z_G = \sum_{n_1} \langle n_1 | e^{-\beta(E_1 n_1 - \mu n_1)} | n_1 \rangle \ldots \sum_{n_\infty} \langle n_\infty | e^{-\beta(E_\infty n_\infty - \mu n_\infty)} | n_\infty \rangle \tag{21.3.34}$$

or

$$Z_G = \prod_k^\infty \mathrm{Tr}_k \ e^{-\beta(E_k n_k - \mu n_k)} \tag{21.3.35}$$

where Tr_k means a trace over the subspace corresponding to the one-particle states $|k\rangle$.

Up to here, the treatment did not depend on whether we were dealing with bosons or fermions. We must now treat the two cases separately.

1) Bosons: Since for bosons the occupation numbers n_k are unrestricted we must sum n_k over all integers. Thus, we get

$$Z_G = \prod_k^\infty \sum_{n=0}^\infty e^{-\beta(E_k - \mu)n} = \prod_k^\infty \left(1 - e^{-\beta(E_k - \mu)}\right)^{-1} \ . \tag{21.3.36}$$

2) Fermions: For fermions the occupation numbers n_k can take on only the values 0 and 1. Thus, we get

$$Z_G = \prod_k^{\infty} \sum_{n=0,1} e^{-\beta(E_k-\mu)n} = \prod_k^{\infty} \left(1 + e^{-\beta(E_k-\mu)}\right) \quad . \tag{21.3.37}$$

If we take the logarithm of Z_G we can treat both systems (bosons and fermions) simultaneously since we have

$$\ln Z_G = \pm \sum_k \ln\left(1 \pm e^{-\beta(E_k-\mu)}\right) \quad . \tag{21.3.38}$$

From now on the upper sign always refers to fermions and the lower sign always refers to bosons.

Introducing the fugacity (equation (21.2.19)) we get

$$\ln Z_G = \pm \sum_k \ln\left(1 \pm z\, e^{-\beta E_k}\right) \quad . \tag{21.3.39}$$

The average number of particles $N = \langle\langle \mathsf{N} \rangle\rangle$ is then, as before, given by

$$N = z \frac{\partial}{\partial z} \ln Z_G \tag{21.3.40}$$

and yields

$$N = \sum_k \frac{z\, e^{-\beta E_k}}{1 \pm z\, e^{-\beta E_k}} = \sum_k \left[e^{\beta(E_k-\mu)} \pm 1\right]^{-1} \quad . \tag{21.3.41}$$

Similarly, we find that the internal energy (when Z_G is expressed in terms of E_k and z) is given by

$$U = -\frac{\partial}{\partial \beta} \ln Z_G \tag{21.3.42}$$

yields

$$U = \sum_k \frac{E_k z\, e^{-\beta E_k}}{1 \pm z\, e^{-\beta E_k}} = \sum_k E_k \left[e^{\beta(E_k-\mu)} \pm 1\right]^{-1} \quad . \tag{21.3.43}$$

To proceed further we must evaluate the sums $\sum_k$. This is most easily accomplished by taking the so-called *thermodynamic limit.* This simply means that we count the density of states (number of modes) in the range k_x and $k_x + dk_x$, k_y and $k_y + dk_y$ as well as k_z and $k_z + dk_z$. This result as already found in chapter 1 yields

$$\frac{V}{(2\pi)^3}\, d^3k \; .$$

One then takes the limit $V \to \infty$. Thus, the sums go over into integrals; in fact

$$\sum_k \to \frac{V}{(2\pi)^3} \int d^3k = \frac{4\pi V}{(2\pi)^3} \int k^2\, dk \quad . \tag{21.3.44}$$

If we now further use that for non-interacting particles

$$E_k = \frac{\hbar^2 k^2}{2m} \tag{21.3.45}$$

and write simply E instead of E_k we get

$$\sum_k \to \frac{V}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \int E^{1/2}\, dE \ . \tag{21.3.46}$$

We can apply these results immediately to the ideal Fermi gas (later we shall also apply them to the ideal Bose gas) and obtain for the average number of particles N and the internal energy U

$$N = \frac{V}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\infty \frac{E^{1/2}\, dE}{z^{-1}e^{\beta E} + 1} \tag{21.3.47}$$

and

$$U = \frac{V}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\infty \frac{E^{3/2}\, dE}{z^{-1}e^{\beta E} + 1} \ . \tag{21.3.48}$$

The next step is to invert equation (21.3.47) and solve for the fugacity z in terms of the mean density of particles

$$\frac{1}{v} = \frac{N}{V} \ . \tag{21.3.49}$$

This result is then substituted into equation (21.3.48) to express the internal energy per unit volume U/V as a function of the mean density $1/v$ and the temperature $1/\beta$.

Unfortunately, even for the simple case of an ideal Fermi gas it is not possible to express the integrals (21.3.47) and (21.3.48) in terms of elementary functions. It is therefore usual to define certain functions in terms of these integrals. Thus, putting

$$\beta E = x^2 \tag{21.3.50}$$

and introducing the *thermal wavelength*

$$\lambda = \left(\frac{2\pi\hbar^2}{mk_BT}\right)^{1/2} \tag{21.3.51}$$

we get

$$\frac{1}{v} = \lambda^{-3}\frac{4}{\sqrt{\pi}}\int_0^\infty \frac{x^2\, dx}{z^{-1}\, e^{x^2} + 1} \tag{21.3.52}$$

and

$$\frac{U}{V} = \beta^{-1}\lambda^{-3}\frac{4}{\sqrt{\pi}}\int_0^\infty \frac{x^4\, dx}{z^{-1}\, e^{x^2} + 1} \ . \tag{21.3.53}$$

We now introduce the functions

$$f_{5/2}(z) = \frac{4}{\sqrt{\pi}} \int_0^\infty x^2\, dx \ln\left(1 + z\, e^{-x^2}\right) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} z^n}{n^{5/2}} \tag{21.3.54}$$

and

$$f_{3/2}(z) = z \frac{\partial}{\partial z} f_{5/2}(z) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} z^n}{n^{3/2}} \ . \tag{21.3.55}$$

Differentiating (21.3.54) under the integral sign shows that

$$f_{3/2}(z) = \frac{4}{\sqrt{\pi}} \int_0^\infty \frac{x^2\, dx}{z^{-1}\, e^{x^2} + 1} \ . \tag{21.3.56}$$

Also, an integration by parts of $f_{5/2}(z)$ shows that

$$f_{5/2}(z) = \frac{2}{3} \frac{4}{\sqrt{\pi}} \int_0^\infty \frac{x^4\, dx}{z^{-1}\, e^{x^2} + 1} \ . \tag{21.3.57}$$

Thus, we have

$$\frac{1}{v} = \lambda^{-3/2} f_{3/2}(z) \tag{21.3.58}$$

$$\frac{U}{V} = \frac{3}{2} \frac{k_B T}{\lambda^3} f_{5/2}(z) \ . \tag{21.3.59}$$

To obtain the corresponding results for an ideal Bose gas requires some care in replacing the sum $\sum_k$ by an integral. This is due to the fact that as $z \to 1$ the single term due to $k = 0$ diverges and may be as important as the entire sum. We therefore first remove the term corresponding to $k = 0$ and replace the rest of the sum by an integral. So, we get for the Bose gas

$$N = \frac{z}{1-z} + \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\infty \frac{E^{1/2}\, dE}{z^{-1} e^{\beta E} - 1} \tag{21.3.60}$$

and

$$U = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\infty \frac{E^{3/2}\, dE}{z^{-1} e^{\beta E} - 1} \ . \tag{21.3.61}$$

Again introducing

$$x^2 = \beta E \quad , \quad \frac{1}{v} = \frac{N}{V} \quad , \quad \lambda = \left(\frac{2\pi\hbar^2}{m k_B T}\right)^{1/2}$$

we get

$$\frac{1}{v} = \frac{1}{V} \frac{z}{1-z} + \lambda^{-3} g_{3/2}(z) \tag{21.3.62}$$

and

$$\frac{U}{V} = \frac{3}{2}\frac{k_B T}{\lambda^3} g_{5/2}(z) \tag{21.3.63}$$

where

$$g_{5/2}(z) = -\frac{4}{\sqrt{\pi}}\int_0^\infty x^2\,dx \ln\left(1 - z\,e^{-x^2}\right) = \sum_{n=1}^{\infty}\frac{z^n}{n^{5/2}} \tag{21.3.64}$$

and

$$g_{3/2}(z) = z\frac{\partial}{\partial z} g_{5/2}(z) = \sum_{n=1}^{\infty}\frac{z^n}{n^{3/2}} \ . \tag{21.3.65}$$

To invert the function $f_{3/2}(z)$ or $g_{3/2}(z)$ to express z in terms of v leads to very complicated expressions. Since our purpose was only to introduce the techniques of statistical mechanics we refer the reader, for further discussions, to one of the standard texts listed at the end of this chapter.

21.4 General Properties of the Density Matrix

The use of the density matrix extends beyond the confines of statistical mechanics to all systems for which it is desirable to have a description of the extent to which the states are specified. Thus even pure, or completely specified states can be described by means of the density matrix. It is desirable to have criteria on the density matrix itself such that one can distinguish density matrices for pure and impure or incompletely specified systems. We develop such criteria as well as other formal properties of the density matrix in this section.

As our starting point we take the defining equation (21.2.9) for the density matrix

$$\rho = \sum_n |n\rangle p_n \langle n| \tag{21.4.66}$$

where the p_n are classical probabilities so that

$$\sum_n p_n = 1 \ . \tag{21.4.67}$$

With this condition we immediately obtain the normalization

$$\begin{aligned} \mathrm{Tr}\rho &= \sum_{m,n}\langle m|n\rangle p_n \langle n|m\rangle \\ &= \sum_{m,n}\delta_{m,n} p_n \\ &= \sum_n p_n = 1 \end{aligned} \tag{21.4.68}$$

so that

$$\mathrm{Tr}\rho = 1 \ . \tag{21.4.69}$$

It is also immediately clear that ρ is self-adjoint

$$\rho^\dagger = \rho \, . \tag{21.4.70}$$

The diagonal matrix elements of ρ are

$$\langle m|\rho|m\rangle = p_m \tag{21.4.71}$$

and are clearly real and satisfy $0 \leq p_m \leq 1$.

Consider a density matrix for a pure state. In this case the probability for the pure state say $|m\rangle$ is $p_m = 1$. All other states have zero probability. Thus,

$$p_n = \delta_{n,m} \tag{21.4.72}$$

so

$$\rho = \sum_n |n\rangle\delta_{n,m}\langle n| = |m\rangle\langle m| \ . \tag{21.4.73}$$

Clearly

$$\rho^2 = \rho \, . \tag{21.4.74}$$

Thus, if ρ describes a pure state it is necessarily idempotent i.e. it satisfies equation (21.4.74). We now show that the converse is also true. Assume ρ is idempotent, that is, equation (21.4.74) is valid. We then have

$$\begin{aligned}\sum_n |n\rangle p_n\langle n| &= \sum_m |m\rangle p_m\langle m|\sum_n |n\rangle p_n\langle n| \\ &= \sum_{n,m} |m\rangle p_m p_n \delta_{m,n}\langle n| \ . \end{aligned} \tag{21.4.75}$$

Hence,

$$\sum_n |n\rangle p_n\langle n| = \sum_n |n\rangle p_n^2\langle n| \ . \tag{21.4.76}$$

It therefore follows that

$$p_n^2 = p_n \tag{21.4.77}$$

so that

$$p_n = 0, 1 \quad \text{for all} \quad n \ . \tag{21.4.78}$$

But

$$\sum_n p_n = 1. \tag{21.4.79}$$

It therefore follows that only one of the probabilities $p_n = 1$ and all others are zero. Thus, ρ describes a pure state.

We have so far been working in a diagonal representation for ρ. This is a direct consequence of the assumption of random phases for the wavefunctions of the external world, i.e. $\overline{\langle c_n | c_m \rangle}$ for $n \neq m$. One can, however, start from a different viewpoint. In that case the density matrix is not automatically diagonal, but nevertheless self-adjoint, so that it can still be brought to diagonal form by a unitary transformation. In the representation in which ρ is diagonal, one can again interpret the wavefunctions for the external world as satisfying the assumption of random phases.

Now suppose we can write the Hamiltonian for our system as well as the external world as

$$H = H_0 + H_I + H_{external} \tag{21.4.80}$$

where H_0 is the Hamiltonian for our system, $H_{external}$ is the Hamiltonian for the external world and H_I is the coupling of our system to the external world. If we now consider the external world as providing only a heat bath then, in fact, we perform time averages over the wavefunctions of the external world. Thus, we are led to consider

$$H = H_0 + \tilde{H}_I + \langle H_{external} \rangle \tag{21.4.81}$$

where now $\langle H_{external} \rangle$ is a constant and $\tilde{H}_I$ acts as an external interaction (potential) on our system which still has the internal dynamics H_0. We can now drop all reference to the external world and ignore the constant $\langle H_{external} \rangle$. Under the action of this Hamiltonian, the density matrix will evolve according to the evolution operator.

$$U(t, t_0) = \exp\left[-\frac{i}{\hbar}(H_0 + \tilde{H}_I)(t - t_0)\right] \tag{21.4.82}$$

so that

$$\rho(t) = U(t, t_0)\rho(t_0)U^\dagger(t, t_0) \ . \tag{21.4.83}$$

Thus, the equation of motion for the density operator is

$$i\hbar\frac{\partial \rho}{\partial t} = [H, \rho] \, . \tag{21.4.84}$$

This equation is known as the *Liouville equation* and, although similar to the Heisenberg equations of motion, clearly has the commutator reversed. Furthermore, we are working in the Schrödinger picture, that is, the wavefunctions carry all the time dependence and the operators are time independent. If we transform to the Heisenberg picture and the Hamiltonian H is independent of time then ρ becomes a constant (time independent) operator. Equation (21.4.84) represents the quantum mechanical version of Liouville's theorem.

This concludes our formal treatment of the density operator. In the next section we apply some of these results to a system with finite degrees of freedom.

21.5 The Density Matrix and Polarization

As stated before, the density matrix is particularly useful in a discussion of systems for which the states are not pure or completely specified. This occurs in the case of scattering of particles with spin when the particles are only partially (or not at all) polarized. The density matrix is well suited to a discussion of this case, as well as the case of complete polarization.

To illustrate this use of the density matrix, we apply it to a discussion of spin 1/2 particles. For a discussion of arbitrary spins, the reader is referred to the literature cited at the end of this chapter [21.1].

If we consider a beam of spin 1/2 particles, say electrons, and if we are only interested in the spin-orientation or polarization of this beam, then we have a system with two degrees of freedom. Thus, the system is completely specified by two complex numbers a, b i.e. the wavefunction can be written

$$\chi = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{21.5.85}$$

where

$$|a|^2 + |b|^2 = 1 \,. \tag{21.5.86}$$

We now describe this system in terms of a density matrix. This is possible since the physical system is completely specified by the *three* components of polarization

$$\mathbf{p} = \langle \vec{\sigma} \rangle \tag{21.5.87}$$

and also involves only *three* real parameters. A general 2×2 hermitean matrix can be written

$$\rho = u_0 1 + \mathbf{u} \cdot \vec{\sigma} \tag{21.5.88}$$

where $(u_0, \mathbf{u})$ are real parameters. Requiring that

$$\mathrm{Tr}\rho = 1 = 2\mathrm{u}_0 \tag{21.5.89}$$

leaves us with only three independent real parameters. These are related to the polarization through (21.5.87).

$$\mathbf{p} = \langle \vec{\sigma} \rangle = \mathrm{Tr}(\rho\vec{\sigma}) = \mathrm{Tr}[\mathrm{u}_0\vec{\sigma} + \vec{\sigma}(\mathbf{u} \cdot \vec{\sigma})] = 2\mathbf{u} \,. \tag{21.5.90}$$

Thus,

$$\rho = \frac{1}{2}[1 + \mathbf{p} \cdot \vec{\sigma}] \,. \tag{21.5.91}$$

To relate ρ directly to χ we consider the matrix

$$\rho' = \chi\chi^\dagger = \begin{pmatrix} a \\ b \end{pmatrix} (a^*, b^*) = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix} \,. \tag{21.5.92}$$

Clearly,

$$\mathrm{Tr}\rho' = |a|^2 + |b|^2 \, . \tag{21.5.93}$$

From (21.5.85) we get

$$\langle\vec{\sigma}\rangle = (\chi, \vec{\sigma}\chi) = \mathbf{p} = (2\,\Re(a^*b), 2\,\Im(a^*b), |a|^2 - |b|^2) \, . \tag{21.5.94}$$

But,

$$\mathrm{Tr}(\rho'\vec{\sigma}) = (2\,\Re(a^*b), 2\,\Im(a^*b), |a|^2 - |b|^2) \, . \tag{21.5.95}$$

Thus,

$$\rho' = \rho \, . \tag{21.5.96}$$

In this form it is easy to see that

$$\rho\chi = \chi \tag{21.5.97}$$

and

$$\rho^2 = \rho \tag{21.5.98}$$

so that ρ corresponds to the pure state with polarization $\mathbf{p}$. This is seen from

$$\vec{\sigma} \cdot \mathbf{p}\,\chi = \chi \, . \tag{21.5.99}$$

If our beam of particles passes through a magnetic field $\mathbf{B}$, the polarization will be changed due to the connection between magnetic moment and spin. The corresponding Hamiltonian[1] is

$$H = -\vec{\mu} \cdot \mathbf{B} = -\frac{1}{2}\frac{ge\hbar}{mc}\vec{\sigma} \cdot \mathbf{B} \, . \tag{21.5.100}$$

More generally the Hamiltonian would simply be an hermitean 2 × 2 matrix

$$H = \frac{1}{2}(A_0 1 + \mathbf{A} \cdot \vec{\sigma}) \tag{21.5.101}$$

with $(A_0, \mathbf{A})$ constants. We then get

$$\begin{aligned} i\hbar\frac{d\mathbf{p}}{dt} &= i\hbar\frac{d}{dt}\mathrm{Tr}(\rho\vec{\sigma}) \\ &= i\hbar\mathrm{Tr}\left(\frac{\partial\rho}{\partial t}\vec{\sigma}\right) \\ &= \mathrm{Tr}([H, \rho]\vec{\sigma}) \\ &= \mathrm{Tr}(H\rho\vec{\sigma} - \rho H\vec{\sigma}) \\ &= \mathrm{Tr}(\rho[\vec{\sigma}, H]) \end{aligned} \tag{21.5.102}$$

---

[1] Actually H is not an Hamiltonian in the sense of the total energy of the system since there is no kinetic energy term. Nevertheless, H is the generator of time translation for the "observable" σ.

where we have used (21.4.84) and the cyclic property of the trace. Applying this result to the specific Hamiltonian (21.5.101) we get:

$$\begin{aligned} i\hbar\frac{d\mathbf{p}}{dt} &= \frac{1}{2}\mathrm{Tr}(\rho[\vec{\sigma},\mathbf{A}\cdot\vec{\sigma}]) \\ &= \frac{1}{2}\mathrm{Tr}(\rho i(\mathbf{A}\times\vec{\sigma})) \\ &= i\mathbf{A}\times\mathbf{p} \end{aligned} \tag{21.5.103}$$

so that

$$\hbar\frac{d\mathbf{p}}{dt} = \mathbf{A}\times\mathbf{p}\,. \tag{21.5.104}$$

This equation is similar to the Euler equation for a symmetrical top in classical mechanics. Furthermore, a magnetic field can only rotate the spin. This is easily seen by considering

$$\frac{d}{dt}\mathbf{p}^2 = 2\mathbf{p}\cdot\frac{d\mathbf{p}}{dt} = \frac{2}{\hbar}\mathbf{p}\cdot(\mathbf{A}\times\mathbf{p}) = 0\,. \tag{21.5.105}$$

The polarization vector is thus seen to maintain its length and is only rotated.

For cases of spin greater than 1/2 as occurs for certain nuclei, the discussion must be generalized. If the spin is j, then the density matrix will be an hermitean $(2j+1)$ by $(2j+1)$ matrix with unit trace. Thus, it requires more than the 3 components of the polarization vector to specify the density matrix. For spin 1, (see problem 22.6) the density matrix can be determined in terms of the 3 components of the polarization vector $\mathbf{p}$ and the 5 components of the quadrupole polarization tensor Q_{ij}.

21.6 Composite Systems

The main purpose of this section is to develop the density matrix formalism for composite systems and to show under what circumstances the density matrix of the composite system is determined by the density matrices of the individual systems.

We consider two systems S_1 and S_2 with respectively k and j degrees of freedom and corresponding coordinates $q_1,\ldots,q_k$ and $x_1,\ldots,x_j$. The composite system $S_c = S_1\otimes S_2$ therefore has $n = k+j$ degrees of freedom. The inner products in the corresponding Hilbert spaces are

$$(\phi,\psi)_1 = \int \phi^*(q_i)\psi(q_i)\,dq_1\ldots dq_k \tag{21.6.106}$$

$$(\phi,\psi)_2 = \int \phi^*(x_i)\psi(x_i)\,dx_1\ldots dx_j \tag{21.6.107}$$

and

$$(\phi,\psi)_c = \int \phi^*(q_i,x_l)\psi(q_i,x_l)\,dq_1\ldots dq_k\,dx_1\ldots dx_j\;\;. \tag{21.6.108}$$

The corresponding observables are labelled $A^{(1)}$, $A^{(2)}$ and $A^{(c)}$. Any observable $A^{(1)}$ in S_1 is naturally also an observable in S_c. The same is true for any $A^{(2)}$ in S_2. If $\{|1,m\rangle\}$, and $\{|2,n\rangle\}$ form bases in the hilbert spaces of S_1 and S_2 respectively, then,

$$|c;n,m\rangle = |1,m\rangle|2,n\rangle \tag{21.6.109}$$

forms a basis for the hilbert space of S_c.

The matrix correspondence between $A^{(1)}$ considered as an element of S_1 or S_c is now given by

$$\begin{aligned}\langle c;m,n|A^{(1)}|c;m',n'\rangle &= \langle 1,m|A^{(1)}|1,m'\rangle\langle 2,n|2,n'\rangle \\ &= \langle 1,m|A^{(1)}|1,m'\rangle\delta_{n,n'} \ .\end{aligned} \tag{21.6.110}$$

The correspondence for $A^{(2)}$ is analogous.

We now consider density matrices in S_1, S_2 and $S_1 \otimes S_2$ labelled $\rho^{(1)}$, $\rho^{(2)}$ and $\rho^{(c)}$ respectively. Now any density matrix $\rho^{(c)}$ in S_c determines a density matrix $\rho^{(i)}$ in S_i $(i=1,2)$. To see how this correspondence is made we consider matrix elements of the density matrices referred to the bases we have given. Then using $\langle\langle\,\cdot\,\rangle\rangle$ to indicate statistical averages we have:

$$\langle\langle A^{(i)}\rangle\rangle = \mathrm{Tr}\left(\rho^{(i)}A^{(i)}\right) = \sum_{m,m'}\rho^{(i)}_{m,m'}A^{(i)}_{m,m'} \tag{21.6.111}$$

for the averages evaluated in S_i. Evaluated in S_c we get

$$\begin{aligned}\langle\langle A^{(1)}\rangle\rangle &= \mathrm{Tr}\left(\rho^{(c)}A^{(1)}\right) \\ &= \sum_{m,m';n,n'} \rho^{(c)}_{m,n;m',n'}A^{(1)}_{m.m'}\delta_{n,n'} \\ &= \sum_{m,m'}\left(\sum_n \rho^{(c)}_{m,n;m',n}\right)A^{(1)}_{m.m'} \ .\end{aligned} \tag{21.6.112}$$

So we have

$$\rho^{(1)}_{m,m'} = \sum_n \rho^{(c)}_{m,n;m',n} \tag{21.6.113}$$

and analogously

$$\rho^{(2)}_{n,n'} = \sum_m \rho^{(c)}_{m,n;m,n'} \ . \tag{21.6.114}$$

Thus, given a density matrix in the composite system determines *uniquely* a density matrix in the subsystem. We now consider the converse question. Under what circumstances do the density matrices $\rho^{(1)}$ and $\rho^{(2)}$ determine the density matrix $\rho^{(c)}$ uniquely? That there is always at least one solution is clear, for if we define

$$\tilde{\rho}_{m,n;m',n'} = \rho^{(1)}_{m,m'}\rho^{(2)}_{n,n'} \tag{21.6.115}$$

then clearly (see problem 22.8) $\tilde{\rho}$ satisfies all the conditions of a normalized density matrix and furthermore (21.6.113) and (21.6.114) are also satisfied. This solution is, however, not unique except when either $\rho^{(1)}$ or $\rho^{(2)}$ corresponds to a pure state (see problem 22.9). Since there are no correlations between S_1 and S_2 in the solution (21.6.115), this is not at all unexpected. For, if this solution were unique, we could never have statistical correlations between subsystems of a system. Of course, if one of the subsystems is a pure state then all the correlations are trivial and (21.6.115) is unique. This is easy to prove. Let $\rho^{(1)}$ correspond to a pure state. Then using the same basis as before we can write without loss of generality

$$\rho^{(1)} = \sum_{m,m'} |1,m\rangle \delta_{1,m}\delta_{1,m'}\langle 1,m'| \,. \tag{21.6.116}$$

Then, we get from (21.6.113)

$$\sum_n \rho^{(c)}_{m,n;m',n} = \delta_{1,m}\delta_{1,m'} \,. \tag{21.6.117}$$

Also, because $\rho^{(c)}$ is a non-negative operator $\rho^{(c)} \geq 0$ we have for m or $m' \neq 1$ that

$$\rho^{(c)}_{m,n;m',n} = 0 \quad \text{for} \quad m \ \text{or} \ m' \neq 1 \,. \tag{21.6.118}$$

Therefore, if $m = m' = 1$ we can write

$$\rho^{(c)}_{1,n;1,n'} = \sum_m \rho^{(c)}_{m,n;m,n'} = \rho^{(2)}_{n,n'} \tag{21.6.119}$$

where we have used (21.6.114). Thus, all matrix elements of $\rho^{(c)}$ are uniquely determined and hence $\rho^{(c)}$ is unique. To summarize we have the following theorem.

Theorem

A density matrix $\rho^{(c)}$ corresponding to a composite system $S_1 \otimes S_2$ is uniquely determined by

$$\tilde{\rho}_{m,n;m',n'} = \rho^{(1)}_{m,m'}\rho^{(2)}_{n,n'} \tag{21.6.120}$$

if and only if either $\rho^{(1)}$ or $\rho^{(2)}$ corresponds to a pure system. In this case $\rho^{(1)}$ and $\rho^{(2)}$ are called the projections of $\rho^{(c)}$ in S_1 and S_2 respectively. The "only if" part of this theorem is proved in problem 22.9. We use this theorem in the next section.

21.7 von Neumann's Theory of Measurement

As an interesting application of the density matrix formalism we give a brief discussion of von Neumann's theory of measurement [21.5] as elaborated by F.W. London and E. Bauer [21.7]. This is not strictly within the bounds of the Copenhagen interpretation. In the viewpoint adopted by N. Bohr, (the strict

Copenhagen interpretation) a formal quantum theory of measurement is not required. This is a consequence of Bohr's belief that all measurements must reduce to classical concepts which are not themselves further reducible since they constitute the ultimate data of sense experience.

The approach taken by von Neumann is closer to the approach we sketched in chapter 7, and will now be described in more detail.

A state $|\Psi\rangle$ evolves in general in two distinctly different ways. If H is the total Hamiltonian of the system then $|\Psi\rangle$ evolves according to the time-dependent Schrödinger equation.

$$i\hbar\frac{\partial}{\partial t}|\Psi\rangle = H|\Psi\rangle \tag{21.7.121}$$

or equivalently

$$|\Psi(t)\rangle = e^{-iHt/\hbar}|\Psi(0)\rangle \ . \tag{21.7.122}$$

This evolution is purely causal. Corresponding to this time development, the density matrix evolves according to

$$\rho(t) = e^{-iHt/\hbar}\,\rho(0)\,e^{iHt/\hbar} \ . \tag{21.7.123}$$

The evolution (21.7.122) can be written in the density matrix formalism by choosing

$$\rho(0) = |\Psi(0)\rangle\langle\Psi(0)| \tag{21.7.124}$$

$$\rho(t) = |\Psi(t)\rangle\langle\Psi(t)| \ . \tag{21.7.125}$$

On the other hand during a measurement process, $|\Psi(0)\rangle$ will change discontinuously and non-causally. Thus if $\{|k\rangle\}$ is a complete set of eigenkets corresponding to the eigenvalues of the observable measured, then during the measurement process, $\rho(0)$ goes over into.

$$\tilde{\rho}(t) = \sum_k |k\rangle|\langle\Psi(0)|k\rangle|^2\langle k| \ . \tag{21.7.126}$$

The two processes

$$\rho(0) \to \rho(t) \tag{21.7.127}$$

and

$$\rho(0) \to \tilde{\rho}(t) \tag{21.7.128}$$

correspond to fundamentally different situations. It is important to note that the density matrix $\tilde{\rho}(t)$ is diagonal; $\tilde{\rho}(t)$ has no off-diagonal elements and so is *incoherent* or classical.

A measurement can always be considered to involve a "system", an "apparatus" and an "observer". According to von Neumann the measurement is completed when the "observer" has made a certain subjective observation. To quote from von Neumann, "Indeed experience only makes statements of this type: an observer has made a certain (subjective) observation; and never like this: a physical quantity has a certain value." This viewpoint is taken in order to avoid an infinite regression. For without this viewpoint we could always attempt to analyze the measurement process further. For example, we might decide to stop when light reflected from the dial of the apparatus hits the retina of the observer's eye. On the other hand we might wish to consider the eye as part of the apparatus and stop with the electrical signal from the optic nerve reaching the brain. Indeed we could go on and consider the chemical changes occurring in the observer's brain and on and on. With the assumption made above, we can stop with any point called the observer. However, to justify this viewpoint it is necessary to show that it does not matter at which stage we stop the analysis. Regardless of how much of the inner workings of the observer are included with the "apparatus" we must always be able to obtain the same result for a measurement as long as it terminates with the subjective awareness of the observer.

The proof of this possibility consists of showing that we can always lump either the "system" and "apparatus" or else the "apparatus" and "observer" together and obtain equivalent results. Before proving these results we dispose of another possibility - whether the statistical character of measurements can depend on the state of knowledge of the observer.

The statistical character of a measurement cannot be due to the observer's lack of knowledge of his (or his apparatus') initial state, since the probability of obtaining a result corresponding to a state $|n\rangle$ is given by $|\langle\Psi|n\rangle|^2$ and is *completely* determined by the state $|\Psi\rangle$. Thus it does *not* depend on the state of the apparatus or the observer. We therefore assume that an observer making an (ideal) measurement is completely aware of his initial state. This simply means that he knows that he is seeing a dial pointing at a given place. Furthermore the apparatus (dial) is also in a pure state.

If apparatus and observer are lumped then the evolution is according to (21.7.126) and the probability of observing the state $|k\rangle$ is $|\langle\Psi|k\rangle|^2$. If, however, we lump the system and apparatus, then evolution according to (21.7.126) should occur only when the observer intervenes.

Thus, we consider a composite system with the original system in an unknown state given by

$$\rho^{(s)} = |\Psi\rangle\langle\Psi| \ . \tag{21.7.129}$$

The apparatus is in a known pure state given by

$$\rho^{(a)} = |a\rangle\langle a| \ . \tag{21.7.130}$$

Let $\{|a_n\rangle\}$ be a complete set of eigenstates of an observable A corresponding to dial readings $\{a_n\}$. The possible eigenstates of the original system corresponding to the observable S being measured are $|k\rangle$, with eigenvalues s_k. The numbering is such that the eigenvalue s_n corresponds to the dial reading a_n. Thus, the composite system (apparatus + system) is in a state given by

$$\rho^{(c)} = \rho^{(s)} \otimes \rho^{(a)} = |\Psi\rangle|a\rangle\langle a|\langle\Psi| \ . \tag{21.7.131}$$

Now the measurement (by the apparatus on the system) corresponds to a unitary evolution as given by (21.7.123)

$$\rho^{(c)}(t) = e^{-iHt/\hbar}\,\rho^{(c)}(0)\,e^{iHt/\hbar}\,. \tag{21.7.132}$$

Thus, according to the observer, a measurement is only performed if he measures the eigenvalues of the simultaneously measurable observables S and A. These pairs of variables s_m, a_n have a probability 0 for $m \neq n$ and $|\langle\Psi|n\rangle|^2$ for $m = n$. This last requirement is dictated by quantum mechanics. In this case the evolution is of the type given by (21.7.126).

$$\rho^{(c)}(t) \to \tilde{\rho}^{(c)}(t) = \sum_k |a_k\rangle|k\rangle|\langle\Psi|k\rangle|^2\langle k|\langle a_k|\,. \tag{21.7.133}$$

If all this holds then the measuring process, so far as it occurs in the apparatus, is explained because the split: system,(apparatus + observer) can also be viewed as the split: (system + apparatus), observer. [2] Thus, our problem reduces to the following: Given a basis set $\{|k\rangle\}$ for our system, find a basis set $\{|a_n\rangle\}$ and state $|a\rangle$ for the apparatus, together with an Hamiltonian H of the form

$$H = H_{system} + H_{apparatus} + H_{interaction} \tag{21.7.134}$$

and a time interval t such that the following holds. If $|\Phi\rangle$ is an arbitrary state of the system and

$$|\Phi', a'\rangle = e^{-iHt/\hbar}\,|\Phi\rangle|a\rangle\ . \tag{21.7.135}$$

Then we have

$$|\Phi', a'\rangle = \sum_m \langle m|\Phi\rangle|m\rangle|a_m\rangle\,. \tag{21.7.136}$$

Notice, that instead of finding an Hamiltonian H, it is sufficient to find an unitary operator U that gives this evolution, since Stone's Theorem (Chapter 6) then guarantees the existence of H. For convenience in proving the existence of U we assume (without loss of generality)

$$|a\rangle = |a_0\rangle\,. \tag{21.7.137}$$

We now *define* the obviously unitary operator U by

$$U \sum_{m,n=-\infty}^{\infty} f_{m,n}|m\rangle|a_m\rangle = \sum_{m,n=-\infty}^{\infty} f_{m,n}|m\rangle|a_{m+n}\rangle\,. \tag{21.7.138}$$

---

[2] In modern terminology the word "observer" is replaced by "environment" and the evolution $\rho^{(c)}(t) \to \tilde{\rho}^{(c)}(t)$ is seen as due to the environment. In either case, observer or environment, this involves a coupling to a many-body system and a loss of coherence i.e. decoherence.

But,

$$U|\Phi\rangle|a\rangle = U\sum_{m=-\infty}^{\infty}\langle m|\Phi\rangle|m\rangle|a_0\rangle$$

$$= \sum_{m=-\infty}^{\infty}\langle m|\Phi\rangle|m\rangle|a_m\rangle\ . \qquad (21.7.139)$$

This completes the proof and yields a completely consistent quantum theory of measurement.

21.8 Decoherence

The approach of von Neumann, although consistent leaves one with an uneasy feeling. It does not explain how the collapse of the wavefunction comes about. To this end it is desirable to find a mechanism that makes the density matrix for a macroscopic observable such as the pointer on a dial become diagonal very rapidly due to its very large number of internal degrees of freedom. This is what is meant by the term decoherence: the coherent parts of the density matrix which are responsible for possible interference effects, the off-diagonal elements, decay rapidly to zero. What remains, the diagonal part, has classical probabilities associated with it and therfore there is no need to bring in the consciousness of the observer. The "consciousness of the observer" in von Neumann's theory is replaced by the very large number of internal degrees of freedom of the macroscopic apparatus.

At present there are only models that show how this comes about in specific cases. There is no general result that demonstrates how decoherence is achieved for an arbitrary macroscopic system. That such a result will be very difficult to find is demonstrated by the fact that certain macroscopic systems, such as superfluids, remain quantum mechanical.

To illustrate the concept we analyze a model of how a macroscopic oscillator in a superposition of two states becomes "classical" when coupled to a very large number of very closely spaced oscillators (the internal vibrations). These internal vibrations may be thought to consist of the phonons (sound vibrations) in the spring of the macroscopic oscillator, but a definite physical system is not specified by this model which is merely defined by its Hamiltonian. The model Hamiltonian is

$$H = \hbar\omega\, a^\dagger a + \sum_k \hbar\omega_k\, b_k^\dagger b_k + \sum_k \left(\lambda_k\, a^\dagger b_k + \lambda_k^*\, a b_k^\dagger\right) \qquad (21.8.140)$$

where the annihilation and creation operators a , $a^\dagger$, b_k , $b_k^\dagger$ satisfy the usual commutation relations and the a , $a^\dagger$ commute with the b_k , $b_k^\dagger$.

As we saw in section 11.5, the best approximation for a "classical" state is a coherent state. Thus, we look for the evolution of a density matrix corresponding to the initial superposition of two coherent states. These individual, normalized, coherent states are labelled $|\alpha, \{\beta_k\}\rangle$ where

$$a\,|\alpha, \{\beta_k\}\rangle = \alpha\,|\alpha, \{\beta_k\}\rangle \tag{21.8.141}$$

$$b_j\,|\alpha, \{\beta_k\}\rangle = \beta_j\,|\alpha, \{\beta_k\}\rangle\ . \tag{21.8.142}$$

The initial state is a superposition of $|\alpha_1(0), \{\beta_k = 0\}\rangle$ and $|\alpha_2(0), \{\beta_k = 0\}\rangle$.

$$|\Psi(0)\rangle = A\,|\alpha_1(0), \{\beta_k = 0\}\rangle + B\,|\alpha_2(0), \{\beta_k = 0\}\rangle\ . \tag{21.8.143}$$

Normalization requires that

$$|A|^2 + |B|^2 = 1\ . \tag{21.8.144}$$

The initial density matrix is

$$\rho(0) = |\Psi(0)\rangle\langle\Psi(0)|\ . \tag{21.8.145}$$

We are interested in the *reduced* density matrix corresponding to tracing over the "environmental" degrees of freedom represented by the $\{\beta_k\}$. The result is

$$\begin{aligned}\rho_r(0) \ &= \ |A|^2\,|\alpha_1(0)\rangle\langle\alpha_1(0)| + |B|^2\,|\alpha_2(0)\rangle\langle\alpha_2(0)| \\ &+ \ AB^*\,|\alpha_1(0)\rangle\langle\alpha_2(0)| + A^*B\,|\alpha_2(0)\rangle\langle\alpha_1(0)|\ .\end{aligned} \tag{21.8.146}$$

This initial density matrix clearly has a non-diagonal part that will produce interference and is non-classical. We now consider the evolution of the density matrix and find the reduced density matrix at some later time t. To do this we differentiate the explicit form of the coherent states (equation (11.6.117)) and assume that the parameters α and β_k are time-dependent. Then we use our model Hamiltonian and write out the Schrödinger equation. Thus, we first evaluate the state $|\Psi(t)\rangle$ and compute the density matrix at time t from this rather than using the Liouville equation.

From (11.6.117) we have

$$|\alpha, \{\beta_k\}\rangle = e^{-|\alpha|^2/2}\,e^{\alpha a^\dagger}\prod_k e^{-|\beta_k|^2/2}\,e^{\beta_k b_k^\dagger}\,|0, \{0\}\rangle\ . \tag{21.8.147}$$

So that, if we assume that the moduli $|\alpha|$ and $|\beta_k|$ remain constant in time, we find that

$$i\hbar\frac{\partial}{\partial t}|\alpha, \{\beta_k\}\rangle = i\hbar\left(a^\dagger\frac{\partial\alpha}{\partial t} + \sum_k b_k^\dagger\frac{\partial\beta_k}{\partial t}\right)|\alpha, \{\beta_k\}\rangle\ . \tag{21.8.148}$$

On the other hand we have, using (21.8.140) - (21.8.142), that

$$H|\alpha, \{\beta_k\}\rangle =$$

$$\left(\hbar\omega a^\dagger\alpha + \sum_k \hbar\omega_k b_k^\dagger\beta_k + \sum_k(\lambda_k a^\dagger\beta_k + \lambda_k^* b_k^\dagger\alpha)\right)|\alpha, \{\beta_k\}\rangle\ . \tag{21.8.149}$$

Combining these results we find that the Schrödinger equation is satisfied if

$$\begin{aligned} i\frac{d\alpha}{dt} &= \omega\alpha + \sum \lambda_k \beta_k \\ i\frac{d\beta_k}{dt} &= \omega_k \beta_k + \lambda_k^* \alpha \,. \end{aligned} \tag{21.8.150}$$

This system of equations is known as the Weisskopf-Wigner system [21.14] and was used to study the effective damping of atomic lines due to interaction with an heat bath. To solve this system of equations we first set

$$\begin{aligned} \alpha(t) &= \hat{\alpha}(t)\, e^{-i\omega t} \\ \beta_k(t) &= \hat{\beta}_k(t)\, e^{-i\omega_k t} \,. \end{aligned} \tag{21.8.151}$$

Then,

$$\begin{aligned} i\frac{d\hat{\alpha}}{dt} &= \sum \lambda_k\, e^{i(\omega-\omega_k)t}\, \hat{\beta}_k \\ i\frac{d\hat{\beta}_k}{dt} &= \lambda_k^*\, e^{-i(\omega-\omega_k)t}\, \hat{\alpha} \,. \end{aligned} \tag{21.8.152}$$

From the equation for $\hat{\beta}_k$ we get

$$\hat{\beta}_k = -i\lambda_k^* \int_0^t e^{-i(\omega-\omega_k)t'}\, \hat{\alpha}(t')\, dt' \ . \tag{21.8.153}$$

After substituting this into the equation for $\hat{\alpha}$ we find

$$i\frac{d\hat{\alpha}}{dt} = -i \sum_k |\lambda_k|^2 \int_0^t e^{i(\omega-\omega_k)(t-t')}\, \hat{\alpha}(t')\, dt' \,. \tag{21.8.154}$$

At this stage we take a "thermodynamic limit" in the sense that we assume that the heat bath has a continuous distribution of oscillators described by a density $n(\omega)$. Then we can replace the sum over k by an integral over ω with the weight $n(\omega)$. Thus, we get

$$\frac{d\hat{\alpha}}{dt} = -\int_0^{\omega_D} |\lambda_{\omega'}|^2 n(\omega')\, d\omega' \int_0^t e^{i(\omega-\omega')(t-t')}\, \hat{\alpha}(t')\, dt' \,. \tag{21.8.155}$$

Here ω_D repesents the *Debye* frequency (see chapter 1). However, to simplify the computation we extend the integration over ω to ∞ and include a small damping factor $e^{-\epsilon(t'-t)}$ in the integrant over t' and then take the limit $\epsilon \to 0$ at the end of the calculation. The resulting expression becomes

$$\frac{d\hat{\alpha}}{dt} = -\int_0^{\infty} |\lambda_{\omega'}|^2 n(\omega')\, d\omega' \int_0^t e^{i(\omega-\omega'+i\epsilon)(t-t')}\, \hat{\alpha}(t')\, dt' \,. \tag{21.8.156}$$

After integrating over t' by parts we get

$$\begin{aligned} \frac{d\hat{\alpha}}{dt} = &- i\int_0^{\infty} |\lambda_{\omega'}|^2 n(\omega')\, d\omega' \left[\frac{1}{\omega-\omega'+i\epsilon}\, \hat{\alpha}(t) \right. \\ &- e^{-i(\omega-\omega')t}\, \frac{1}{\omega-\omega'+i\epsilon}\, \hat{\alpha}(0) \\ &- \left. \int_0^t e^{i(\omega-\omega')(t-t')}\, \frac{1}{\omega-\omega'+i\epsilon}\, \frac{d\hat{\alpha}(t')}{dt'}\, dt' \right] \,. \end{aligned} \tag{21.8.157}$$

Next we examine these terms a little more closely to make some physically justified approximations. In the term on the second line we see that for large times t (large compared to $1/\omega_D$) we get very rapid oscillations which interfere destructively. Thus, we drop these fluctuations. The third term is also small because for weak damping $\hat{\alpha}(t')$ varies slowly and thus has a small derivative. In addition, we get destructive interference except for $t \approx t'$. Thus, we also drop this term. The result is

$$\frac{d\hat{\alpha}}{dt} = -i \int_0^\infty |\lambda_{\omega'}|^2 n(\omega')\, d\omega' \frac{1}{\omega - \omega' + i\epsilon} \hat{\alpha}(t) \ . \tag{21.8.158}$$

So we finally find

$$\frac{d\hat{\alpha}}{dt} = (-i\Delta\omega - \gamma)\hat{\alpha}(t) \tag{21.8.159}$$

where we have used

$$\frac{1}{\omega - \omega' + i\epsilon} = P\frac{1}{\omega - \omega'} - i\pi\delta(\omega - \omega') \tag{21.8.160}$$

to obtain

$$\begin{aligned} \gamma &= \pi|\lambda_\omega|^2 n(\omega) \\ \Delta\omega &= P\int_0^\infty |\lambda_{\omega'}|^2 n(\omega') \frac{1}{\omega - \omega'}\, d\omega' \ . \end{aligned} \tag{21.8.161}$$

Therefore we have that

$$\begin{aligned} \hat{\alpha}(t) &= \alpha(0)\, e^{-i\Delta\omega t - \gamma t} \\ &+ \text{ fluctuations that we have neglected} \end{aligned} \tag{21.8.162}$$

and

$$\begin{aligned} \alpha(t) &= \alpha(0)\, e^{-i(\omega + \Delta\omega)t - \gamma t} \\ &+ \text{ fluctuations that we have neglected} \ . \end{aligned} \tag{21.8.163}$$

This result can now be substituted back into (21.8.153) to obtain $\hat{\beta}_k$ Therefore, we finally find

$$\beta_k(t) = \frac{\lambda_k^* \left[e^{-i(\Delta\omega - i\gamma)t} - 1\right]}{\omega + \Delta\omega - \omega_k - i\gamma} \alpha(0) \ . \tag{21.8.164}$$

Now we can examine the density matrix at time t. Actually all we are interested in is again the reduced density matrix obtained by tracing over the internal degrees of freedom. But first we split the density matrix into two parts: the diagonal or incoherent part ρ_d and the off-diagonal or coherent part ρ_c. The results are

$$\begin{aligned} \rho_d(t) &= A^*A|\alpha_1(t), \{\beta_k(t)\}\rangle\langle\alpha_1(t), \{\beta_k(t)\}| \\ &+ B^*B|\alpha_2(t), \{\beta_k(t)\}\rangle\langle\alpha_2(t), \{\beta_k(t)\}| \end{aligned} \tag{21.8.165}$$

and

$$\begin{aligned}\rho_c(t) &= B^*A|\alpha_1(t),\{\beta_k(t)\}\rangle\langle\alpha_2(t),\{\beta_k(t)\}| \\ &+ A^*B|\alpha_1(t),\{\beta_k(t)\}\rangle\langle\alpha_2(t),\{\beta_k(t)\}|\ . \end{aligned} \tag{21.8.166}$$

Tracing over the diagonal part is trivial and produces the reduced diagonal density matrix

$$\rho_{d,r}(t) = A^*A|\alpha_1(t)\rangle\langle\alpha_1(t)| + B^*B|\alpha_2(t)\rangle\langle\alpha_2(t)|\ . \tag{21.8.167}$$

The coherent part requires a little more calculation.

$$\begin{aligned} &\rho_{c,r}(t) \\ &= B^*A|\alpha_1(t)\rangle\sum_k\langle\{\beta'_k(t)\}|\{\beta_k(t)\}\rangle\langle\alpha_2(t)|\langle\{\beta_k(t)\}|\{\beta'_k(t)\}\rangle \\ &+ A^*B|\alpha_2(t)\rangle\sum_k\langle\{\beta'_k(t)\}|\{\beta_k(t)\}\rangle\langle\alpha_1(t)|\langle\{\beta_k(t)\}|\{\beta'_k(t)\}\rangle . \end{aligned} \tag{21.8.168}$$

To evaluate this we use the fact that

$$\begin{aligned}\langle\{\beta'_k(t)\}|\{\beta_k(t)\}\rangle &= \prod_k\langle\beta'_k(t)|\beta_k(t)\rangle \\ &= \exp\left(-\sum_k\frac{1}{2}|\beta'_k(t)-\beta_k(t)|^2\right)\ . \end{aligned} \tag{21.8.169}$$

Here we have used equation (11.4.77). Next we use the conservation equation (see problem 21.10)

$$\frac{d}{dt}\left(\alpha^*\alpha' + \sum_k\beta^*_k\beta'_k\right) = 0 \tag{21.8.170}$$

to simplify this result to get

$$\prod_k\langle\beta'_k(t)|\beta_k(t)\rangle = \exp\left[-\frac{1}{2}\left(|\alpha_1(0)-\alpha_2(0)|^2 - |\alpha_1(t)-\alpha_2(t)|^2\right)\right]. \tag{21.8.171}$$

After substituting for $\alpha_1(t)$ and $\alpha_2(t)$ we get

$$\prod_k\langle\beta'_k(t)|\beta_k(t)\rangle = \exp\left[-\frac{1}{2}\left(|\alpha_1(0)-\alpha_2(0)|^2(1-e^{-2\gamma t})\right)\right]\ . \tag{21.8.172}$$

The result for the reduced coherent density matrix is

$$\begin{aligned}\rho_{c,r} &= (AB^*|\alpha_1(0)\rangle\langle\alpha_2(0)| + A^*B|\alpha_2(0)\rangle\langle\alpha_1(0)|) \\ &\times \exp\left[-\frac{1}{2}\left(|\alpha_1(0)-\alpha_2(0)|^2(1-e^{-2\gamma t})\right)\right]\ . \end{aligned} \tag{21.8.173}$$

If we assume that initially the macroscopic oscillator was in a superposition of states with momentum zero and positions $x_1(0)$ and $x_2(0)$ then we can use the equation

$$\alpha = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x_0 + ip_0) \tag{21.8.174}$$

to get

$$\begin{aligned} \rho_{c,r} &= (AB^*|\alpha_1(0)\rangle\langle\alpha_2(0)| + A^*B|\alpha_2(0)\rangle\langle\alpha_1(0)|) \\ &\times \exp\left[-\frac{1}{4}\frac{m\omega}{\hbar}(x_1(0) - x_2(0))^2(1 - e^{-2\gamma t})\right] . \end{aligned} \tag{21.8.175}$$

This is the result we wanted to obtain.

To see what this means we have to look at the numerical value of some of the parameters for realistic situations. For this purpose we consider the damping time γ^{-1}. Suppose this damping γ is very small so that the damping time is quite large. Then for short times t we can write $e^{-2\gamma t} \approx 1 - 2\gamma t$. In this case the coherent part of the reduced density matrix is exponentially damped with a damping time τ given by

$$\tau = \frac{4\hbar}{m\omega(x_1(0) - x_2(0))^2} \times \frac{1}{2\gamma} \tag{21.8.176}$$

To see how remarkably small this number turns out to be we choose γ^{-1} to be of the order of one hour, ω of the order of one radian/second, m to be one gram and the separation of the initial positions to be 10^{-6} meter. In this case we find that

$$\tau \approx 1.5 \times 10^{-11} \text{ s} \tag{21.8.177}$$

This means that such a macroscopic quantity becomes classical in about 15 picoseconds.

21.9 Conclusion

We have now completed a development of nonrelativistic quantum mechanics. The theory is consistent and even complete. Nevertheless, from its inception attempts have been made to show that quantum mechanics is only a superficial theory based on a more fundamental hidden variables theory. In the last few years, however, experiments designed to test Bell's inequality [21.9] have forced proponents of hidden variable theories to retreat further and further. It thus seems very unlikely that modifications of the quantum concept will be required in the near future.

There still remains the problem of providing a successful union of quantum mechanics and relativity theory. The groundwork for this was laid in the early thirties. Although the product of this unification, modern quantum field theory, has been highly successful we cannot yet claim to know that quantum field theory is a completely consistent theory. This statement is made in spite of the great advances of the last several decades.

On the more philosophical level some difficult epistemological questions remain. In this book we have, for the most part, employed what is essentially the Copenhagen interpretation of quantum mechanics. There are at least two rival interpretations: the Statistical [21.10] and the Many Worlds [21.11] interpretations. They are radically different in the pictures of reality that they present, yet they agree on all experimental predictions. Thus they are, in principle, experimentally indistinguishable from the Copenhagen interpretation. At this stage it is a matter of personal preference which interpretation is employed. This means that a very large element of subjectivity has entered physics at a very fundamental level.

21.10 Problems

21.1 Compute the average energy of an assembly of identical simple harmonic oscillators using:

a) the microcanonical ensemble
b) the canonical ensemble
c) the grand canonical ensemble.

21.2 Prove the following properties of a density matrix.

a) $\rho^2 \leq \rho$. This implies $\rho \geq 0$.
b) $\mathrm{Tr}([\rho, A]) = 0$

and discuss the sufficient conditions on the operator A for this equation to hold. To see that this is not trivial, consider $\mathrm{Tr}([x, p])$.

21.3 In a gas of electrons, a fraction p are known to have their z-component of spin in the up direction. Assume the remainder are random.

a) What is the average value of s_x, s_y, and s_z?
b) If nothing is known about the spins of the remaining fraction $1 - p$ of electrons what are the maximum possible values of $\langle\langle s_x \rangle\rangle$, $\langle\langle s_y \rangle\rangle$ and $\langle\langle s_x \rangle\rangle$?

21.4 Verify equation (21.2.24).

$$N = \langle\langle \mathsf{N} \rangle\rangle = z \frac{\partial}{\partial z} \ln Z_G \, .$$

21.5 Show that for spin 1, the density matrix can be completely specified by the polarization vector $\mathbf{p}$, and the quadrupole polarization tensor Q_{ik} defined for spin j by

$$\mathbf{p} = \frac{\langle \mathbf{J} \rangle}{j\hbar}$$

$$Q_{ik} = \frac{\langle J_i J_k + J_k J_i \rangle}{j(j+1)\hbar^2} - \frac{2}{3}\delta_{ik} \, .$$

21.6 Show that for the case of general spin j, if we again define the polarization vector by

$$\mathbf{p} = \frac{\langle \mathbf{J} \rangle}{j\hbar}$$

and are given a "Hamiltonian"

$$H = -\gamma \mathbf{J} \cdot \mathbf{B}$$

where **B** is a magnetic field, then we have

$$\frac{d\mathbf{p}}{dt} = \gamma \mathbf{p} \times \mathbf{B}$$

$$\frac{d^2\mathbf{p}}{dt^2} = 0 \; .$$

21.7 Show that $\tilde{\rho}$ defined by (21.6.115) satisfies the general properties of a density matrix

a)

$$\tilde{\rho} = \tilde{\rho}^\dagger$$

b)

$$\mathrm{Tr}\tilde{\rho} = 1$$

c)

$$\tilde{\rho}^2 \leq \tilde{\rho}$$

as well as

d) the equations (21.6.113) and (21.6.114).

21.8 Show that if we have

$$\rho^{(1)} = \alpha_1 \sigma^{(1)} + \beta_1 \tau^{(1)} \;\; , \quad \alpha_1 + \beta_1 = 1 \quad , \quad \alpha_1 \, , \, \beta_1 > 0$$

$$\rho^{(2)} = \alpha_2 \sigma^{(2)} + \beta_2 \tau^{(2)} \;\; , \quad \alpha_2 + \beta_2 = 1 \quad , \quad \alpha_2 \, , \, \beta_2 > 0$$

and

$$\mathrm{Tr}\sigma^{(i)} = \mathrm{Tr}\tau^{(i)} = 1 \; .$$

Then any combination

$$\tilde{\rho} = \alpha \sigma^{(1)} \otimes \sigma^{(2)} + \beta \sigma^{(1)} \otimes \tau^{(2)} + \gamma \tau^{(1)} \otimes \sigma^{(2)} + \delta \tau^{(1)} \otimes \tau^{(2)}$$

with

$$\alpha+\beta=\alpha_1 \quad \alpha+\gamma=\alpha_2$$

$$\gamma+\delta=\beta_1 \quad \beta+\delta=\beta_2$$

satisfies (21.6.113) and (21.6.114) and is a possible density matrix for the composite system. This establishes the necessity of the condition that $\rho^{(1)}$ and $\rho^{(2)}$ correspond to pure states in order that (21.6.115) give a unique solution for a density matrix for the composite system.

21.9 Consider a "gas" of Bose particles with energy either $+E$ or $-E$. The Hamiltonian for this system is

$$H = E(a_2^\dagger a_2 - a_1^\dagger a_1)$$

where a_1 , $a_i^\dagger$ $(i=1,2)$ are the usual annihilation and creation operators for bosons.
a) Show that the canonical partition function is given by

$$Z_N = \frac{\sinh \beta E(N+1)}{\sinh \beta E}$$

and that the grand canonical partition function is given by

$$Z_G = \left[1 - 2z\cosh \beta E + z^2\right]^{-1} \ .$$

b) Compute the internal energy U, and the average number of particles $\langle\langle \mathsf{N} \rangle\rangle$ and express U as a function of β and $\langle\langle \mathsf{N} \rangle\rangle$ rather than as a function of β and z.

21.10 Verify the conservation law (equation (21.8.170))

$$\frac{d}{dt}\left(\alpha^*\alpha' + \sum_k \beta_k^*\beta_k'\right) = 0 \ .$$

21.11 Consider two point masses connected by a spring so that the Hamiltonian is

$$H = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + \frac{1}{2}k\,(x_1 - x_2)^2 \ .$$

Find the density matrix for the nth excited state

$$\rho_n(x_1, x_1'; x_2, x_2') = \psi_{K,n}(x_1, x_2)\psi_{K,n}(x_1', x_2')^*$$

where

$$H\psi_{K,n}(x_1,x_2) = \left[\frac{\hbar^2K^2}{2(m_1+m_2)} + (n+1/2)\hbar\omega\right]\psi_{K,n}(x_1,x_2)$$

and

$$\omega = \sqrt{k/m} \quad , \quad m = \frac{m_1m_2}{m_1+m_2} \ .$$

Finally compute the reduced density matrix for particle 1 and show that it is not idempotent; i.e. that correlations between the two particles persist.

Bibliography

[21.1] A rather lengthy review article, treating many different aspects of the density matrix with numerous references is:
D. Ter Haar, *Theory and Applications of the Density Matrix*, Reports on Progress in Physics **24**, 304-362 (1961).

[21.2] A very enjoyable little book on statistical thermodynamics with a viewpoint different from ours is:
E. Schrödinger, *Statistical Thermodynamics*, Cambridge University Press (1962).

Two of the more standard texts on statistical mechanics are:

[21.3] L.D. Landau and E.M. Lifshitz, *Statistical Physics*, Addison-Wesley Publishing Co., Inc., U.S.A. (1958)

[21.4] K. Huang, *Statistical Mechanics*, John Wiley and Sons, Inc., New York (1963).

[21.5] One of the original books treating the density matrix and discussing the quantum theory of measurement is:
J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press (1955).

[21.6] Further theories of measurement with references to the original papers are to be found in:
M. Jammer, *The Philosophy of Quantum Mechanics*, John Wiley and Sons, Inc., New York (1974).

[21.7] F. London and E. Bauer, *La Théorie de l'Observation en Mécanique Quantique*, Hermann & Cie., Paris (1939)

[21.8] For a different approach to polarization see:
W.H. McMaster, Polarization and the Stokes Parameters, Amer. J. Phys. **22**, 351 (1954).

[21.9] S.J. Freedman and J.F. Clauser, Experimental test of local hidden-variable theories, Phys. Rev. Lett. **28**, 938-941 (1972)

[21.10] L.E. Ballentine, The Statistical Interpretation of Quantum Mechanics, Rev. Mod. Phys. **42**, 358-381 (1970).

[21.11] Bryce S. De Witt and Neill Graham, *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton University Press (1973).

The idea of decoherence is discussed at length in the following two references.

[21.12] P. Busch, P.J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement*, Springer Verlag (1996).

[21.13] R. Omnes, *The Interpretation of Quantum Mechanics*, Princeton University Press (1994)

[21.14] V.Weisskopf and E. Wigner, Zeitschrift Phys. **63**, 54 (1930).

Index

NONRELATIVISTIC QUANTUM MECHANICS

Third Edition

The main unique feature of *Nonrelativistic Quantum Mechanics* is its discussion of Hilbert space and rigged Hilbert space. This invaluable book is suitable for advanced undergraduate students as well as graduate students

Reviews of the first edition:

"This book should be carefully considered for use in advanced undergraduate or early graduate courses."

Nature

"The overriding merit of Capri's book is that he exploits the requisite mathematics, without fuss, at a level appropriate to an introductory course, taking great pains to explain its relevance to the physics ... the book is already a must for teachers, and selected chapters from it a godsend for intelligent and self-respecting students."

Contemporary Physics

"We have enjoyed very much reviewing quantum mechanics, as presented in this excellent textbook. Anton Capri must be complimented for succeeding in the difficult task of striving to be as rigorous as possible mathematically without leading the reader to wander away from direct relevancy to physics."

Foundations of Physics

"Careful, comprehensive and comprehensible, this book will be appreciated as a valuable reference for students and teachers — something to turn to when you can't remember what Franck and Hertz did, or how to get the Lippmann–Schwinger equations."

Physics Bulletin

ISBN 981-02-4651-X (pbk)

9 789810 246518

www.worldscientific.com
4712 sc