# LECTURES ON AN INTRODUCTION TO QUANTUM MECHANICS

J. Tinsley Oden

 $\begin{array}{c} {\rm ICES} \\ {\rm The \ University \ of \ Texas \ at \ Austin} \\ 2006 \end{array}$ 

## Preface

I began preparing these notes during the summer of 2004 as a bases for weekly lectures I gave to a group of colleagues: friends, students, post docs, and professional associates. The goal of the course, taught and taken completely voluntarily, was simply to learn some basic things about quantum mechanics and its relationship to molecular dynamics. In preparing these notes, I relied on several excellent textbooks, treatises, and monographs on the subject. This monograph can be regarded as a second edition, in which I corrected typos in the first edition and added material on spin, quantum chemistry, DFT, and related topics. We found the subject beautiful and exciting, and that it is quite accessible to those with backgrounds in classical mechanics, applied mathematics, partial differential equations, and functional analysis. Hopefully, with further reading and exposure, the theory will become a tool in our work in computational engineering and sciences.

J. Tinsley Oden 2006

# Contents

1	Elec	tromagnetic Waves	1
	1.1	Electrical Fields	1
	1.2	Gauss' Law	4
	1.3	Electric Potential Energy	5
	1.4	Magnetic Fields	6
	1.5	Some Properties of Waves	8
	1.6	Maxwell's Equations	12
	1.7	Electromagnetic Waves	16
<b>2</b>	Intr	oduction to Quantum Mechanics	17
	2.1	Introductory Comments	17
	2.2	The Photoelectric Effect	18
	2.3	The Compton Effect	19
	2.4	Heisenberg's Uncertainty Principle	19
	2.5	The Correspondence Principle	21
	2.6	Schrödinger's Equation	21
	2.7	Elementary Properties of the Wave Equation	25
	2.8	Momentum	26
	2.9	Wave Packets and Fourier Transforms	27
	2.10	The Wave-Momentum Duality	28
	2.11	Appendix	29
3	Dyn	amical Variables and Observables in Quantum Mechanics:	
	The	Mathematical Formalism	33
	3.1	Introductory Remarks	33
	3.2	The Hilbert Space $L^2(\mathbb{R})$ (or $L^2(\mathbb{R}^N)$ )	34
	3.3	Dynamical Variables and Hermitian Operators	35

#### CONTENTS

	$3.4 \\ 3.5 \\ 3.6 \\ 3.7$	Spectral Theory of Hermitian OperatorsObservables and Statistical DistributionsThe Continuous SpectrumThe Generalized Uncertainty Principle	37 40 41 42		
4	App	olications: The Harmonic Oscillator and the Hydrogen Atom	45		
	4.1	Introductory Remarks	45		
	4.2	Ground States and Energy Quanta: the Harmonic Oscillator	45		
	4.3	The Hydrogen Atom	47		
<b>5</b>	Spir	and the Pauli Principle	55		
	5.1	Angular Momentum and Spin	55		
	5.2	Identical Particles and the Pauli Principle	58		
	5.3	The Helium Atom	60		
	5.4	Variational Principle	63		
	5.5	Summary of Atomic Structure: The Periodic Table	64		
6	The	Transition from Quantum Mechanics to Approximate The-			
	orie	s and to Molecular Dynamics	69		
	6.1	Introduction	69		
	6.2	Molecular Dynamics	69		
	6.3	The Connection Between MD and Quantum Mechanics	72		
	6.4	Some Approximate Methods	74		
		6.4.1 The LCAO Method	74		
		6.4.2 The Hartree-Fock Model	76		
		6.4.3 Density Functional Theory	78		
Bi	Bibliography				

# List of Figures

1.1	a) A charge $q$ at a point $P$ in the plane, and electric field lines emanating from $P$ and b) the electric field produced by positive	
	charges $q_1$ and $q_2$ at points P and Q	3
1.2	Equal and opposite charges on a line	4
1.3	Equal and opposite charges in an electric field $E$ a distance $d$ apart	4
1.4	Model of an atom as charged electrons in orbits around a nucleus	6
1.5	Electric dipole analogous to a dipole caused by a magnet	8
1.6	Incremental motion of a wave from over a time increment $\Delta t$	10
1.7	Properties of a simple plane wave	11
1.8	Wavelengths and frequencies of electromagnetic waves.	15
1.9	Components of electromagnetic waves	15
2.1	Diffraction of a beam of electrons	20
2.2	Histogram showing the number of people versus age	30
2.3	Two histograms with the same median, average, and probable	
	value, but different standard deviations	32
5.1	The Periodic Table	65
5.2	The simplistic diagrams ignore orbital structures	66
5.3	Figure 5.3	66
5.4	Figure 5.4	67
5.5	Figure 5.5	67
6.1	The hydrogen molecule $H_2$ consisting of two nuclei at $a$ and $b$ and	
	electrons at 1 and 2. $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$	72
6.2	A molecular system consisting of $M$ nuclei and $N$ electrons	77

#### LIST OF FIGURES

Chapter 1

# **Electromagnetic Waves**

#### **1.1** Electrical Fields

The now classical science of electricity and magnetism recognizes that material objects can possess what is called an *electric charge*—an intrinsic characteristic of the fundamental particles that make up the objects. The charges in objects we encounter in everyday life may not be apparent, because the object is *electrically neutral*, carrying equal amounts of two kinds of charges, *positive charge* and *negative charge*. Atoms consist of positively charged *protons*, negatively charged *electrons*, and electrically neutral *neutrons*, the protons and neutrons being packed together in the *nucleus* of the atom.

The mathematical characterization of how charges interact with one another begins with Coulomb's Law, postulated in 1785 by Charles Augustine Coulomb on the basis of experiments. It is stated as follows: consider two charged particles (or point charges) of magnitude  $q_1$  and  $q_2$  separated by a distance r. The electrostatic force of attraction or repulsion between the charges has magnitude

$$F = k \frac{|q_1| |q_2|}{r^2}$$
(1.1)

where k is a constant, normally expressed in the form

$$k = \frac{1}{4\pi\epsilon_0} = 8.99 \times 10^9 \text{ N m}^2/\text{C}^2$$

where  $\epsilon_0$  is the permittivity constant,

$$\epsilon_0 = 8.85 \times 10^{-12} \text{ N m}^2/\text{C}^2$$

Here C is the SI measure of a unit charge, called a Coulomb, and is defined as the amount of charge that is transferred across a material wire in one second due to a 1-ampere current in the wire. The reason for this choice of k is made clear later.

According to Jackson [9], the inverse-square dependence of force on distance (of Coulomb's law) is known to hold over at least 24 orders of magnitude in the length scale.

Two important properties of charge are as follows:

1. charge is quantized: Let e denote the elementary charge of a single electron or proton, known from experiments to be

$$e = 1.60 \times 10^{-19} \text{ C}$$

then any positive or negative charge q is of the form,

$$q = ne$$
, where  $n = \pm 1, \pm 2, \pm 3, \dots$   $(n \in \mathbb{Z})$ 

The fact that electrical charge is "quantized" (meaning discretely defined as an integer multiple of e) is regarded by some as "one of the most profound mysteries of the physical world" (Cf. [9, page 251])

2. charge is *conserved*: the net charge in a system or object is preserved: the net charge is constant unless additional charged particles are added to the system.

The fundamental notion of an electric field is intimately tied to the force field generated by electrical charges. Let  $q_1$  denote a positive point charge situated at a point P in space. Imagine that a second positive point charge  $q_2$  is placed at point Q near to P. According to Coulomb's law,  $q_1$  exerts a repulsive electrostatic force on  $q_2$ . This vector field of forces is called the *electric field*. We say that  $q_1$ sets up an electric field E in the space surrounding it, such that the magnitude of E at a point Q depends upon the distance from P to Q and the direction depends on the direction from P to Q and the electrical sign of  $q_1$  at P.

In practice,  $\boldsymbol{E}$  is determined at a point by evaluating the electrostatic force  $\boldsymbol{F}$  due to a positive test charge  $q_0$  at that point (see Fig. 1.1a). Then  $\boldsymbol{E} = q_0^{-1} \boldsymbol{F}$ .  $\boldsymbol{E}$  thus has the units of Newtons per Coulomb (N/C). The magnitude of the electric field, then, due to a point charge q is

$$q_0^{-1}\left(\frac{1}{4\pi\epsilon_0}\frac{|q_1|\,|q_0|}{r^2}\right) = \frac{|q_1|}{4\pi\epsilon_0 r^2}$$

For m such charges,

$$\boldsymbol{E} = q_0^{-1} \sum_{i=1}^{m} \boldsymbol{F}_{0i}$$
(1.2)



(b) Positive charges at points P and Q

**Figure 1.1**: a) A charge q at a point P in the plane, and electric field lines emanating from P and b) the electric field produced by positive charges  $q_1$  and  $q_2$  at points P and Q.

 $F_{0i}$  being the force from 0, the point of application of  $q_0$ , and the *m* charges  $q_i$ ,  $i = 1, 2, \ldots, m$ . The electric field lines for two positive charges are illustrated in Fig. 1.1b.

A fundamental question arises concerning the electrical force between two point charges  $q_1$  and  $q_2$  separated by a distance r; if  $q_2$  is moved toward  $q_1$ , does the electric field change immediately? The answer is no. The information that one of the charges has moved travels outwardly in all directions at the speed of light c as an electromagnetic wave. More on this later.

The concept of an *electric dipole* is also important. The electric field of two particles of charge q but opposite sign, a distance d apart, at a point x on an axis through the point charges, is

$$\boldsymbol{E}(x) = \frac{q}{4\pi\epsilon_0} \left[ \frac{1}{(x-d/2)^2} - \frac{1}{(x+d/2)^2} \right] \boldsymbol{i}$$



Figure 1.2: Equal and opposite charges on a line



Figure 1.3: Equal and opposite charges in an electric field E a distance d apart

(see Fig. 1.2). For  $x \gg d$ ,

$$\boldsymbol{E}(x) = \frac{1}{2\pi\epsilon_0} \frac{qd}{x^3} \, \boldsymbol{i}$$

The vector

$$\boldsymbol{m} = qd\boldsymbol{i} \tag{1.3}$$

is called the *electric dipole moment*, and

$$\boldsymbol{E}(x) = \boldsymbol{m}/2\pi\epsilon_0 x^3 \tag{1.4}$$

The torque  $\boldsymbol{\tau}$  created by a dipole moment in an electric field  $\boldsymbol{E}$  is  $\boldsymbol{m} \times \boldsymbol{E}$ , as can be deduced from Fig. 1.3

A fundamental property of charged bodies is that charge is *conserved*, as noted earlier. The charge of one body may be passed to another, but the net charge remains the same. No exceptions to this observation have ever been found.

### 1.2 Gauss' Law

A closely related idea is that embodied in *Gauss' Law*, which asserts that the net charge contained in a bounded domain is balanced by the flux of the electric field

#### 1.3. ELECTRIC POTENTIAL ENERGY

through the surface. This is merely a restatement of the property noted earlier that charge is conserved.

The net flux of an electric field  $\boldsymbol{E}$  through the boundary surface  $\partial \Omega$  of a bounded region  $\Omega$  is given by

$$q_{\Omega} = \oint_{\partial \Omega} \epsilon_0 \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}A \tag{1.5}$$

where  $\boldsymbol{n}$  is a unit normal to  $\partial\Omega$  and  $q_{\Omega}$  is the net charge enclosed in the region  $\Omega$  bounded by the surface  $\partial\Omega$ 

The relationship between Gauss' law and Coulomb's law is immediate: Let  $\Omega$  be a sphere of radius r containing a positive point charge q at its center. Then

$$\epsilon_0 \oint_{\partial(\text{sphere})} \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}A = \epsilon_0 E(4\pi r^2) = q \; ,$$

 $\mathbf{SO}$ 

$$E = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \; ,$$

which is precisely Coulomb's law. This relationship is the motivation for choosing the constant  $k = 1/(4\pi\epsilon_0)$  in Coulomb's law.

#### **1.3** Electric Potential Energy

When an electrostatic force acts between two or more charged particles within a system of particles, an electrostatic potential energy U can be assigned to the system such that in any change  $\Delta U$  in U the electrostatic forces do work. The potential energy per unit charge (or due to a charge q) is called the *electric* potential V; V = U/q. V has units of volts, 1 volt = 1 joule/coulomb.

Atom Models. The attractive or repulsive forces of charged particles leads directly to the classical Rutherford model of an atom as a tiny solar system in which electrons, with negative charges e, move in orbits about a positively charged nucleus. Thus, when an electron travels close to a fixed positive charge q = +e, it can escape the pull or reach a stable orbit, spinning around the charge and thereby creating a primitive model of an atom. This primitive model was discarded when Bohr introduced the quantum model of atoms in which, for any atom, electrons



Figure 1.4: Model of an atom as charged electrons in orbits around a nucleus

can only exist in so-called discrete quantum states of well-defined energy or socalled energy shells. The motion of an electron from one shell to another can only happen instantaneously in a quantum jump. In such a jump, there is obviously a change in energy, which is emitted as a proton of electromagnetic radiation (more on this later). Bohr's model was later improved by the probability density model of Schrödinger.

#### 1.4 Magnetic Fields

Just as charged objects produce electric fields, magnets produce a vector field  $\boldsymbol{B}$  called a *magnetic field*. Such fields are created by moving electrically charged particles or as an intrinsic property of elementary particles, such as electrons. In this latter case, magnetic fields are recognized as basic characteristics of particles, along with mass, electric charge, etc. In some materials the magnetic fields of all electrons cancel out, giving no net magnetic field. In other materials, they add together, yielding a magnetic field around the material.

There are no "monopole" analogies of magnetic fields as in the case of electric fields, i.e., there are no "magnetic monopoles" that would lead to the definition of a magnetic field by putting a test charge at rest and measuring the force acting on a particle. Instead, we consider a particle of charge q moving through a poing P with velocity  $\boldsymbol{v}$ . A force  $\boldsymbol{F}_B$  is developed at P. The magnetic field  $\boldsymbol{B}$  at P is defined as the vector field such that

$$\boldsymbol{F}_B = q\boldsymbol{v} \times \boldsymbol{B} \tag{1.6}$$

Its units are *tesla*'s: T = newton/(coulomb)(meter/second) = N/(C/s)(m) or *gauss*'s (10<sup>-4</sup> tesla). Since the motion of electric charge is called *current*, it is

#### 1.4. MAGNETIC FIELDS

easily shown that B can be expressed in terms of the current *i* for various motions of charges. Ampere's Law asserts that on any closed loop  $\mathscr{C}$  in a plane,

$$\oint_{\mathscr{C}} \boldsymbol{B} \cdot \, \mathrm{d}\boldsymbol{s} = \mu_0 i_{\text{enclosed}}$$

where  $\mu_0$  is the *permeability constant*,  $\mu_0 = 4\pi \times 10^{-7}$  T-m/A, and  $i_{\text{enclosed}}$  is the net current flowing perpendicular to the plane in the planar region enclosed by  $\mathscr{C}$ . For motion of a charge along a straight line,  $\boldsymbol{B} = \mu_0 i/2\pi R$ , R being the perpendicular distance from the infinite line.

Ampere's law does not take into account the induced magnetic field due to a change in the electric flux. When this is taken into account, the above equality is replace by the *Ampere-Maxwell Law*,

$$\oint_{\mathscr{C}} \boldsymbol{B} \cdot \, \mathrm{d}\boldsymbol{s} = \mu_0 i_{\text{enclosed}} + \mu_0 \epsilon_0 \frac{\mathrm{d}}{\mathrm{d}t} \Phi_E$$

where  $\Phi_E$  is the electric flux,

$$\Phi_E = \int_A \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}A$$

 $\boldsymbol{n}$  being a unit normal to the area A circumscribed by the closed loop  $\mathscr{C}$ .

Just as the motion of a charged particle produces a magnetic field, so also does the change of a magnetic field produce an electric field. This is called an *induced electrical field* and is characterized by *Faraday's Law*: Consider a particle of charge q moving around a closed loop  $\mathscr{C}$  encompassing an area A with unit normal n. Then B induces an electric field E such that

$$\oint_{\mathscr{C}} \boldsymbol{E} \cdot d\boldsymbol{s} = -\frac{\mathrm{d}}{\mathrm{d}t} \int_{A} \boldsymbol{B} \cdot \boldsymbol{n} \,\mathrm{d}A$$
(1.7)

From the fact that magnetic materials have poles of attraction and repulsion, it can be appreciated that magnetic structure can exist in the form of magnetic dipoles. Magnetic monopoles do not exist. For this reason, the net magnetic flux through a closed Gaussian surface must be zero:

$$\oint_{\partial\Omega} \boldsymbol{B} \cdot \boldsymbol{n} \, \mathrm{d}A = 0 \tag{1.8}$$

This is referred to as Gauss' Law for magnetic fields.

Every electron has an intrinsic angular momentum s, called the *spin angular* momentum, and an intrinsic *spin magnetic dipole moment*  $\mu$ , related by

$$\boldsymbol{\mu} = -\frac{e}{m}\boldsymbol{s} \tag{1.9}$$

where e is the elementary charge  $(1.6 \times 10^{-19} \text{ C})$  and m is the mass of an electron  $(9.11 \times 10^{-31} \text{ kg})$ 

#### CHAPTER 1. ELECTROMAGNETIC WAVES



Figure 1.5: Electric dipole analogous to a dipole caused by a magnet

### 1.5 Some Properties of Waves

The concept of a wave is a familiar one from everyday experiences. In general, a wave is a perturbation or disturbance in some physical quantity that propagates in space over some period of time. Thus, an acoustic wave represents the space and time variation of pressure perturbations responsible for sound; water waves, the motion of the surface of a body of water over a time period; electromagnetic waves, as will be established, characterize the evolution of electrical and magnetic fields over time, but need no media through which to move as they propagate in a perfect vacuum at the speed of light.

Mathematically, we can characterize a wave by simply introducing a function u of position x (or  $\mathbf{x} = (x_1, x_2, x_3)$  in three dimensions) and time t. In general, waves can be represented as the superposition of simple sinusoidal functions, so that the building blocks for wave theory are functions of the form,

$$u(x,t) = u_0 \mathrm{e}^{\mathrm{i}(kx-\omega t)} \tag{1.10}$$

or, for simplicity, of the form

$$u(x,t) = u_0 \sin(kx - \omega t)$$
(1.11)

which are called plane waves. A plot of this last equation is given in Fig. 1.7. Here

We also define

$$\lambda = 2\pi/k = \text{the wave length} T = 2\pi/\omega = \text{the period (of oscillation)}$$
 (1.13)

#### 1.5. SOME PROPERTIES OF WAVES

The frequency v of the wave is defined as

$$v = 1/T \tag{1.14}$$

The wave speed v is defined as the rate at which the wave pattern moves, as indicated in Fig. 1.6. Since point A retains its position on the wave crest as the wave moves from left to right, the quantity

$$\varphi = kx - \omega t$$

must be constant. Thus

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t} = 0 = k\frac{\mathrm{d}x}{\mathrm{d}t} - \omega$$



Figure 1.6: Incremental motion of a wave from over a time increment  $\Delta t$ 

so the wave speed is

$$v = \frac{\omega}{k} = \frac{\lambda}{T} = \lambda\nu$$
(1.15)

The quantity

$$\varphi = kx - \omega t =$$
the phase of the wave (1.16)

Two waves of the form

$$u_1 = u_0 \sin(kx - \omega t)$$
 and  $u_2 = u_0 \sin(kx - \omega t + \varphi)$ 

have the same amplitude, angular frequency, angular wave number, wave length, and period, but are *out-of-phase* by  $\varphi$ . These waves produce *interference* when superimposed:

$$u(x,t) = u_1(x,t) + u_2(x,t) = (2u_0 \cos \varphi/2) \sin(kx - \omega t + \varphi/2)$$
(1.17)



Figure 1.7: Properties of a simple plane wave

for

 $\varphi = 0$ , the amplitude doubles while the wave length and period remains the same (constructive interference)

 $\varphi = \pi$ , the waves cancel  $(u(x, t) \equiv 0)$  (destructive interference). Observe that

$$\frac{\partial^2 u}{\partial t^2} = -\omega^2 u_0 \sin(kx - \omega t)$$

and

$$\frac{\partial^2 u}{\partial x^2} = -k^2 u_0 \sin(kx - \omega t)$$

so that u satisfies the second-order (hyperbolic) wave equation

$$\frac{\partial^2 u}{\partial t^2} - \left(\frac{\omega^2}{k^2}\right) \frac{\partial^2 u}{\partial x^2} = 0 \tag{1.18}$$

and, again,  $\omega/k$  is recognized as the wave speed.

# 1.6 Maxwell's Equations

Consider electric and magnetic fields E and B in a vacuum. Recall the following physical laws:

Gauss' Law (Conservation of charge in a volume  $\Omega$  enclosed by a surface  $\partial \Omega$ )

$$\epsilon_0 \int_{\partial\Omega} \boldsymbol{E} \cdot \boldsymbol{n} \, \mathrm{d}A = q_\Omega = \int_\Omega \rho \, \mathrm{d}x \tag{1.19}$$

where  $\boldsymbol{n}$  is a unit vector normal to the surface area element dA and  $\rho$  is the charge density,  $q_{\Omega}$  being the total charge contained in  $\Omega$ .

Faraday's Law (The induced electromotive force due to a charge in magnetic flux)

$$\int_{\mathscr{C}} \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} = -\frac{\mathrm{d}}{\mathrm{d}t} \int_{A} \boldsymbol{B} \cdot \boldsymbol{n} \, \mathrm{d}A \tag{1.20}$$

where  $\mathscr{C}$  is a closed loop surrounding a surface of area A and  $\boldsymbol{n}$  is a unit normal to dA. The total eletromotive force is  $\int_{\mathscr{C}} \boldsymbol{E} \cdot d\boldsymbol{s}$ .

The Ampere-Maxwell Law (The magnetic field produced by a current i)

$$\int_{\mathscr{C}} \boldsymbol{B} \cdot d\boldsymbol{s} = \mu_0 i_{\text{enclosed}} + \mu_0 \epsilon_0 \frac{d}{dt} \left( \int_A \boldsymbol{E} \cdot \boldsymbol{n} \, dA \right)$$
$$= \mu_0 \int_A \boldsymbol{j} \, dA + \mu_0 \epsilon_0 \int_A \boldsymbol{n} \cdot \frac{\partial \boldsymbol{E}}{\partial t} \, dA \qquad (1.21)$$

#### 1.6. MAXWELL'S EQUATIONS

where  $\mathscr{C}$  is a closed curve surrounding a surface of area A, j is the current density, and n is a unit vector normal to dA.

The Absence of Magnetic Monopoles

$$\int_{\partial\Omega} \boldsymbol{B} \cdot \boldsymbol{n} \, \mathrm{d}A = 0 \tag{1.22}$$

where  $\partial \Omega$  is a surface bounding a bounded region  $\Omega \subset \mathbb{R}^3$ .

Applying the divergence theorem to the left-hand sides of (1.19) and (1.22) and Stokes' theorem to the left-hand sides of (1.20) and (1.21), and arguing that the integrands of the resulting integrals must agree almost everywhere, gives the following system of equations:

$$\epsilon_{0} \nabla \cdot \boldsymbol{E} = 4\pi\rho$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}$$

$$\nabla \times \boldsymbol{B} = \mu_{0} \boldsymbol{j} + \mu_{0} \epsilon_{0} \frac{\partial \boldsymbol{E}}{\partial t}$$

$$\nabla \cdot \boldsymbol{B} = 0$$
(1.23)

These are Maxwell's equations of classical electromagnetics of macroscopic events in a vacuum. In materials other than a vacuum, the permittivity  $\epsilon$  and the magnetic permeability  $\mu$  may be functions of position  $\boldsymbol{x}$  in  $\Omega$ , and must satisfy constitutive equations

$$\begin{aligned} D &= \epsilon_0 E + P &= \text{ the electric displacement field} \\ H &= \mu_0^{-1} B - M &= \text{ the magnetic field} \end{aligned}$$
 (1.24)

with  $\boldsymbol{B}$  now called the magnetic inductance,  $\boldsymbol{P}$  the polarization vector and  $\boldsymbol{M}$  the magnetic dipole. Then the equations are rewritten in terms of  $\boldsymbol{D}$ ,  $\boldsymbol{H}$ ,  $\boldsymbol{B}$ , and  $\boldsymbol{j}$ :

$$\nabla \cdot \boldsymbol{D} = \rho_f$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}$$

$$\nabla \cdot \boldsymbol{B} = 0$$

$$\nabla \times \boldsymbol{H} = \boldsymbol{j}_f + \frac{\partial \boldsymbol{D}}{\partial t}$$
(1.25)

where  $\rho_f$  and  $\boldsymbol{j}_f$  are appropriately scaled charge and current densities. When quantum and relativistic effects are taken into account, an additional term appears in the first and third equation and  $\partial/\partial t$  is replaced by the total material time-derivative,

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} \; ,$$

 $\boldsymbol{v}$  being the velocity of the media.

Let j = 0. Then

$$rac{\partial}{\partial t} oldsymbol{
abla} imes oldsymbol{B} = oldsymbol{
abla} imes rac{\partial oldsymbol{B}}{\partial t} = -oldsymbol{
abla} imes oldsymbol{
abla} imes oldsymbol{B} = \mu_0 \epsilon_0 rac{\partial^2 oldsymbol{E}}{\partial t^2}$$

so that we arrive at the wave equation,

$$\frac{\partial^2 \boldsymbol{E}}{\partial t^2} - \frac{1}{\mu_0 \epsilon_0} \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{E} = \boldsymbol{0}$$
(1.26)

where the wave speed is

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} = 3.0 \times 10^8 \text{ m/s}$$
 (1.27)

which is the speed of light in a vacuum. *Electromagnetic disturbances (electro-magnetic waves) travel (radiate) at the speed of light through a vacuum.* 



Figure 1.8: Wavelengths and frequencies of electromagnetic waves.



Figure 1.9: Components of electromagnetic waves

#### 1.7 Electromagnetic Waves

Consider a disturbance of an electric (or magnetic) field. According to what we have established thus far, this disturbance is radiated as an electromagnetic wave that has the following properties:

• The electrical field component E and the magnetic field component B are normal to the direction of propagation v of the wave:

$$oldsymbol{
u} = oldsymbol{E} imes oldsymbol{B} / |oldsymbol{E} imes oldsymbol{B}|$$

• **E** is normal to **B**:

$$\boldsymbol{E}\cdot\boldsymbol{B}=0$$

- The wave travels at the speed of light in a vacuum
- The fields E and B have the same frequency and are in phase with one another.

Thus, while the wave speed  $c = \omega/k = \lambda/T$  is constant, the wave length  $\lambda$  can vary enormously. The well-known scales of electromagnetic wave lengths is given in the accompanying figure.

Some electromagnetic waves such as X-rays and visible light are radiated from sources that are of atomic or nuclear dimension. The quantum transfer of an electron from one shell to another radiates an electromagnetic wave. The wave propagates an energy packet called a proton in the direction  $\boldsymbol{\nu}$  of propagation. We explore this phenomenon in more detail in the next chapter.

CHAPTER 2	2
-----------	---

# Introduction to Quantum Mechanics

#### 2.1 Introductory Comments

Quantum mechanics emerged as a theory put forth to resolve two fundamental paradoxes that arose in describing physical phenomena using classical physics. First, can physical events be described by waves (optics, wave mechanics) or particles (the mechanics of corpuscle or particles) or both? Second, at atomic scales, experimental evidence confirms that events occur at quantum levels and are not continuous in time. Waves are characterized by frequency  $\nu$  and wave number  $\kappa = k/(2\pi)$  (or period  $T = 1/\nu$  or wave length  $\lambda = 1/\kappa$ ) while the motion a particle is characterized by its total energy E and its momentum p. In atomic physics and in the case of electromagnetic waves, particles carry a definite quanta of energy and momentum and  $\nu \lambda = c$ . There elementary physical phenomena has a dual aspect and can be described by either waves or particles, and the principal physical attributes are related according to

$$E = hv = \hbar\omega$$
 ,  $p = \frac{h}{\lambda} = \hbar k$ 

where h is Planck's constant, and  $\hbar = h/2\pi$ .

Matter Waves. In 1925, de Broglie put forth the hypothesis that the same dualism between the wave and corpuscle theories of light occur in matter; i.e., a material particle will have a matter wave corresponding to it just as light quanta has a light wave. The wave characteristics are related through the formula,  $E = h\nu$  and  $p = h/\lambda$ .

The situation is succinctly put forth by Messiah [13, page 59]:

"...microscopic objects have a very general property: they appear under two separate irreconcilable aspects, the wave aspect, exhibiting the superposition property characteristic of waves, and the corpuscular aspect on the other hand, namely localized grains of energy and momentum. There exists a universal relationship between these two aspects given by

 $E=h\nu \quad,\qquad p=h/\lambda \;"$ 

#### 2.2 The Photoelectric Effect

The photoelectric effect provides one of the first examples of a phenomena that cannot be described by the classical theory of light but can be captured by a corpuscular theory that treats light as a beam of light quanta, photons, of energy  $h\nu$ ,  $\nu$  being the frequency of the radiation.

The experimental setup involves subjecting an alkali metal to short-wavelength light in a high vacuum. The surface becomes positively charged as it gives off negative electricity in the form of electrons. Very precise experiments enable one to measure the total current and the velocity of the electrons. Observations confirm that

- 1. the velocity of the emitted electrons depends only on the frequency  $\nu$  of the light (the electromagnetic wave)
- 2. the emission of electrons is observed immediately at the start of radiation (contrary to classical theory)
- 3. the energy of the electron is, within a constant M of the material, proportional to the frequency  $\nu$ , with constant of proportionality h, Planck's constant:

$$E = h\nu - M$$

This last result is interpreted as follows: every light quantum striking the metal collides with an electron to which it transfers its energy. The electron loses part of its energy M due to the work required to remove it from the metal. The velocity of the electron does not depend on the intensity of the light, but the number of electrons emitted does, and is equal to the number of incident light quanta.

Property 3) is inferred by Einstein's 1905 note on the photoelectric effect in which, in generalization of Planck's theory, he postulated that light radiation consists of a beam of protons of energy  $h\nu$  and velocity c, and he showed how this hypothesis could explain the photoelectric affect. Experiments by Meyer and Gerlach in 1914 were in excellent agreement with this proposition.

#### 2.3 The Compton Effect

The corpuscular nature of light was observed by Compton in 1922 in experiments in which a block of paraffin was bombarded by X-rays. It was observed that radiation scattered at an angle of less than 90° possesses a wave length greater than the primary radiation so that the frequency  $\nu'$  of the scattered wave is smaller than the frequency  $\nu$  of the incident radiation, a phenomena that cannot be reconciled with classical wave theory.

Compton explained the process as one in which a light quantum (a photon) of energy  $h\nu$  strikes an electron, transferring kinetic energy to the electron while losing energy itself. The scattered photon has smaller energy  $h\nu'$ . The Compton formula,

$$\Delta \lambda = 2 \frac{h}{mc} \sin^2 \frac{\theta}{2}$$

gives the change in wave length of the photon due to the scattering process, m being the mass of the electron and  $\theta$  the angle between the incident wave and the direction of the scattered light.

## 2.4 Heisenberg's Uncertainty Principle

We have seen that physical phenomena can be interpreted in terms of corpuscles (particles) or in terms of waves. The uncertainty principle of Heisenberg addresses the surprising fact that it is impossible to determine that it is corpuscles or waves that one observes in a phenomena; more specifically, the position and momentum of an electron cannot be determined simultaneously. Wave and corpuscular views of nature events are said to be in *duality* (or to be complementary) in that if the corpuscular character of an experiment is proved, it is impossible to prove at the same time its wave character. Conversely, proof of a wave characteristic means that the corpuscular characteristic, at the same time, cannot be established.

A classical example illustrating these ideas involves the diffraction of a beam of electrons through a slit, as indicated in Fig. 2.1. As the electrons pass through the slit, of width  $\Delta y$ , the beam is accompanied by diffraction and diverges by an angle  $\theta$ . The electron is represented by a de Broglie wave of wave length  $\lambda = h/p$ . Thus,

$$\Delta y \sin \theta \approx \lambda = h/p$$



Figure 2.1: Diffraction of a beam of electrons

Likewise, the change in momentum in the y-direction is

$$\Delta p = \Delta p_y = p \sin \theta$$

$$\Delta y \Delta p_y \approx h$$
(2.1)

The precise position of an electron in the slit cannot be determined, nor can the variation in the momentum be determined with greater precision that  $\Delta p_y$  (or  $h/\Delta y$ ).

The relation (2.1) is an example of Heisenberg's uncertainty principle. Various other results of a similar structure can be deduced from quantum mechanics. Some examples and remarks follow:

1. A more precise analysis yields

$$\Delta y = \sqrt{\langle y^2 \rangle - \langle y \rangle^2} , \qquad \Delta p = \sqrt{\langle p^2 \rangle - \langle p \rangle^2}$$
$$\Delta y \Delta p \ge \frac{1}{2}h$$

where  $\langle y \rangle$  is the average measurement of position around y and similar definitions apply to p.

Thus,

2.

$$\Delta t \Delta E \ge h$$

Thus, just as position and momentum cannot be localized in time, a change in energy  $\Delta E$  which accompanies a change in frequency and cannot be localized in time.

#### 2.5 The Correspondence Principle

It is difficult to improve upon Born's description of the correspondence principle: "The leading idea (Bohr's correspondence principle, 1923) may be broadly stated as follows. Judged by the test of experience, the laws of classical physics have brilliantly justified themselves in all processes of motion, macroscopic and microscopic, down to the motions of atoms as a whole (kinetic theory of matter). It must therefore by laid down, as an unconditionally necessary postulate, that the new mechanics, supposed still unknown, must in all problems reach the same results as the classical mechanics. In other words, it must be demonstrated that, for the limiting cases of large masses and of orbits of large dimension, the new mechanics passes over into classical mechanics."

The major shortcoming of the classical theory is seen at the microscopic scale of atomistic physics where discontinuities (quanta) appear. Classical theory accounts for phenomena in the limit where discontinuities are infinitely small. "Quantum theory must approach Classical Theory in the limit of large quantum numbers." [13, page 29]

#### 2.6 Schrödinger's Equation

Wave and particle mechanics are brought together in a way inspired by the uncertainty principle by introducing a wave function  $\Psi(\boldsymbol{x}, t)$  associated with a particle from which the probability that the particle can be found at point  $\boldsymbol{x}$  and at time t can be deduced. This wave function completely determines the dynamical state of the quantum system in the sense that all information on the dynamics of the system at time t can be deduced from knowledge of  $\Psi$ . The central goal of quantum theory, then, is this: knowing  $\Psi$  at an initial time  $t_0$ , determine  $\Psi$ at later times t. To accomplish this, we must derive an equation of propagation of the wave  $\Psi$ , and this equation must connect the wave and corpuscular entities in a manner consistent with the observations made earlier. Finally, the resulting propagation equation must be consistent with the correspondence principle. The Case of a Free Particle. Considering first the wave equation of a free particle (ignoring hereafter relativistic effects), we begin with the fact that a wave  $\Psi = \Psi(\boldsymbol{x}, t)$  is a superposition of monochromatic plane waves (here in one space dimension).

As we have seen earlier, plane waves are of the form,

$$\Psi(x,t) = \psi_0 \mathrm{e}^{\mathrm{i}(kx - \omega t)}$$

where  $\psi_0$  is the amplitude. Then general wave function will be a superposition of waves of this form. Since now the wave number k and the angular frequency  $\omega$  are related to energy and momentum according to  $k = 2\pi/\lambda = p/\hbar$  and  $\omega = 2\pi/\nu = E/\hbar$ , we have

$$\Psi(x,t) = \psi_0 \mathrm{e}^{\mathrm{i}(px - Et)/\hbar}$$

Thus,

$$\frac{\partial \Psi}{\partial t} = -\frac{i}{\hbar} E \psi_0 e^{i(px - Et)/\hbar} = -\frac{i}{\hbar} E \Psi$$
$$\frac{\partial \Psi}{\partial x} = \frac{i}{\hbar} p \psi_0 e^{i(px - Et)/\hbar} = \frac{i}{\hbar} p \Psi$$

Thus, E and p can be viewed as operators on  $\Psi$ :

For a free particle of mass m the energy E and momentum p are related by

$$E = \frac{p^2}{2m} \tag{2.3}$$

Thus, introducing (2.2) into (2.3), we arrive at the following partial differential equation for the wave functions:

$$i\hbar\frac{\partial\Psi}{\partial t} + \frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} = 0$$
(2.4)

This is Schrödinger's equation for a free particle.

Superposition in  $\mathbb{R}^n$ . In general, we may use superposition of waves:

$$\Psi(\boldsymbol{x},t) = \int_{\mathbb{R}^n} \varphi(\boldsymbol{p}) \mathrm{e}^{\mathrm{i}(\boldsymbol{p}\cdot\boldsymbol{x}-Et)/\hbar} \mathrm{d}p$$

#### 2.6. SCHRÖDINGER'S EQUATION

to obtain

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{x},t) + \frac{\hbar^2}{2m}\Delta\Psi(\boldsymbol{x},t) = 0$$

 $\Delta =$  Laplacian. What is  $\phi$ ? Let t = 0. Then the initial condition is

$$\Psi(\boldsymbol{x},0) = \varphi(\boldsymbol{x}) = \int_{\mathbb{R}^n} \varphi(\boldsymbol{p}) \mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{x}} \mathrm{d}\boldsymbol{p}$$

i.e.,  $\varphi(\boldsymbol{p})$  is the Fourier transform of the initial data,  $\Psi(\boldsymbol{x}, 0)$ .

Hamiltonian Form. A popular way of writing Schrödinger's equation is to introduce the Hamiltonian,

$$H(p,q) = E \qquad (q \sim x)$$

Then the Hamiltonian operator is written

$$H(p,q) = H\left(\frac{h}{2\pi i}\frac{\partial}{\partial q},q\right)$$

and Schrödinger's equation becomes,

$$\left(H\left(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial q},q\right) + \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial t}\right)\Psi = 0$$
(2.5)

The Case of Potential Energy. If the particle is under the action of a force with potential energy V = V(q), then

$$H(p,q) = \frac{p^2}{2m} + V(q)$$
 (2.6)

and Schrödinger's equation becomes

$$\left(-\frac{1}{2m}\hbar^2\frac{\partial^2}{\partial q^2} + V(q) + \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial t}\right)\Psi = 0$$
(2.7)

Multiple Particles. The number of independent variables in the domain of the wave function, in addition to time t, is equal to the number of particles associated with the atom or atoms under study. Thus, a complex atom consisting of a nucleus at position  $\boldsymbol{x}_0$  and Z electrons at positions  $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_Z$  has a wave function  $\Psi = \Psi(\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_Z, t)$ . We describe the Hamiltonian and the Schrödinger equation for such systems below.

*Relativistic Quantum Mechanics.* While we are not going to cover relativistic effects in these notes, the wave equation for this case easily follows from the fact that in this case, the energy of a free particle is given by

$$E^2 = p^2 c^2 + m^2 c^4$$

From this we deduce the equation

$$\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta + \left(\frac{mc}{\hbar}\right)^2\right]\Psi(\boldsymbol{x},t) = 0$$
(2.8)

This is called the Klein-Gordon equation. The Klein-Gordon extension does not take into account the intrinsic spin of elementary particles such as an electron. Dirac introduced spin as an additional degree of freedom that allowed the square root of  $E^2$  to be computed, resulting in the relativistic quantum mechanic equation called the Dirac equation. In this setting, the wave function is a four-component spinor rather than a scalar-valued function. In non-relativistic quantum mechanics, to which we restrict ourselves, the property of particle spin is introduced as a postulate. We discuss this subject in more detail in Chapter 5.

General Formulations of Schrödinger's Equations. The basic plan is to consider a dynamical system of N particles with coordinates  $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_N$  and momenta  $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_N$  for which the Hamiltonian is a functional,

$$H = H(\boldsymbol{q}_1, \boldsymbol{q}_2, \dots, \boldsymbol{q}_N; \boldsymbol{p}_1, \boldsymbol{p}_2, \dots, \boldsymbol{p}_N; t)$$

To this dynamical system there corresponds a quantum system represented by a wave function  $\Psi(\boldsymbol{q}_1, \boldsymbol{q}_2, \dots, \boldsymbol{q}_N, t)$ . Setting

$$E = i\hbar \frac{\partial}{\partial t}$$
 and  $\mathbf{p}_r = \frac{\hbar}{i} \frac{\partial}{\partial \mathbf{q}_r}$ ;  $r = 1, 2, \dots, N$  (2.9)

Schrödenger's equation becomes,

$$\begin{cases} i\hbar\frac{\partial}{\partial t} - H\left(\boldsymbol{q}_{1},\boldsymbol{q}_{2},\ldots,\boldsymbol{q}_{N};\frac{\hbar}{i}\frac{\partial}{\partial\boldsymbol{q}_{1}},\frac{\hbar}{i}\frac{\partial}{\partial\boldsymbol{q}_{2}},\\ \ldots,\frac{\hbar}{i}\frac{\partial}{\partial\boldsymbol{q}_{N}} \end{cases} \right) \end{cases} \Psi(\boldsymbol{q}_{1},\boldsymbol{q}_{2},\ldots,\boldsymbol{q}_{N},t) = 0$$

$$(2.10)$$

The rules (2.9) showing the correspondence of energy and momentum to the differential operators shown hold only in cartesian coordinates in the form indicated, as Schrödinger's equation should be invariant under a rotation of the coordinate axes.

# 2.7 Elementary Properties of the Wave Equation

We will undertake a basic and introductory study of Schrödinger's equation, first for a single particle in one space dimension, and then generalize the analysis by considering a more general mathematical formalism provided by function space settings. An excellent source for this level of treatment is Griffith's book, *Introduction to Quantum Mechanics* [6], but the books of Born [2] and Messiah [13] should also be consulted.

1. Review. The Schrödinger equations governing the dynamics of a single particle of mass m moving along a line, the x-axis, subjected to forces derived from a potential V is

$$i\hbar\frac{\partial\Psi}{\partial t} + \frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} - V\Psi = 0$$
(2.11)

where

$$\hbar = \frac{h}{2\pi}$$
 = Planck's constant  $/2\pi = 1.054573 \times 10^{-34}$  J s  
 $\Psi = \Psi(x, t)$  = the wave function

To (2.11) we must add an initial condition

$$\Psi(x,0) = \varphi(x) \tag{2.12}$$

where  $\varphi$  is prescribed. We postulate that the wave function has the property

$$\Psi^*\Psi = |\Psi(x,t)|^2 = \rho(x,t)$$

where  $\Psi^*$  = the complex conjugate of  $\Psi(\Psi = \psi_0 e^{-i(kx - \omega t)})$  and

$$\rho(x,t) =$$
 the probability distribution function  
associated with  $\Psi$ , such that  
 $\rho(x,t)dx =$  the probability of finding the particle  
between x and  $x + dx$  at time t

The wave function must be normalized since  $\rho$  is a PDF:

$$\int_{-\infty}^{\infty} |\Psi(x,t)|^2 \,\mathrm{d}x = 1 \tag{2.13}$$

**Postulate 2.1**. If  $\Psi = \Psi(x,t)$  is a solution of Schrödinger's equation (2.11), then

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{-\infty}^{\infty} |\Psi(x,t)|^2 \,\mathrm{d}x = 0 \tag{2.14}$$

Proof.

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{-\infty}^{\infty} |\Psi|^2 \,\mathrm{d}x = \int_{-\infty}^{\infty} \frac{\partial}{\partial} \left( \Psi^*(x,t)\Psi(x,t) \right) \mathrm{d}x$$
$$= \int_{-\infty}^{\infty} \left( \Psi^* \frac{\partial\Psi}{\partial t} + \frac{\partial\Psi^*}{\partial t} \Psi \right) \mathrm{d}x$$

But

$$\begin{array}{lll} \displaystyle \frac{\partial \Psi}{\partial t} & = & \displaystyle \frac{\mathrm{i}\hbar}{2m} \frac{\partial^2 \Psi}{\partial x^2} + \frac{\mathrm{i}}{\hbar} V \Psi \\ \displaystyle \frac{\partial \Psi^*}{\partial t} & = & \displaystyle - \frac{\mathrm{i}\hbar}{2m} \frac{\partial^2 \Psi^*}{\partial x^2} - \frac{\mathrm{i}}{\hbar} V \Psi^* \end{array}$$

 $\operatorname{So}$ 

$$\frac{\partial}{\partial t} |\Psi|^2 = \Psi^* \frac{\partial \Psi}{\partial t} + \frac{\partial \Psi^*}{\partial t} \Psi = \frac{\mathrm{i}\hbar}{2m} \left( \Psi^* \frac{\partial^2 \Psi}{\partial x^2} - \frac{\partial^2 \Psi^*}{\partial x^2} \Psi \right) = \frac{\mathrm{i}\hbar}{2m} \frac{\partial}{\partial x} \left( \Psi^* \frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x} \Psi \right)$$
(2.15)

Hence,

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{-\infty}^{\infty} |\Psi(x,t)|^2 \,\mathrm{d}x = \frac{\mathrm{i}\hbar}{2m} \left( \Psi^* \frac{\Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x} \Psi \right) \Big|_{x \to -a}^{x \to +a}$$
$$= 0$$

since  $\Psi, \Psi^* \to 0$  as  $x \to \pm \infty$  in order that (2.13) can hold.

# 2.8 Momentum

As noted earlier, momentum can be viewed as an operator,

$$p\Psi = \left(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\right)\Psi \tag{2.16}$$

#### 2.9. WAVE PACKETS AND FOURIER TRANSFORMS

Another way to interpret this is as follows. If  $\langle x \rangle$  is the expected value of the position x of the particle,

$$\frac{\mathrm{d}\langle x\rangle}{\mathrm{d}t} = \int_{-\infty}^{\infty} x \frac{\partial}{\partial t} (\Psi^* \Psi) \mathrm{d}t$$

$$= \frac{\mathrm{i}\hbar}{2m} \int_{-\infty}^{\infty} x \frac{\partial}{\partial x} \left( \Psi^* \frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x} \Psi \right) \mathrm{d}x \qquad (\text{from (2.15)})$$

$$= -\frac{\mathrm{i}\hbar}{2m} \int_{-\infty}^{\infty} \left( \Psi^* \frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x} \Psi \right) \mathrm{d}x \qquad (\text{integrating by parts})$$

$$= -\frac{\mathrm{i}\hbar}{m} \int_{-\infty}^{\infty} \Psi^* \frac{\partial \Psi}{\partial x} \mathrm{d}x$$

$$= \frac{1}{m} \int_{-\infty}^{\infty} \Psi^* \left( \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial x} \right) \Psi \mathrm{d}x$$

$$= \frac{1}{m} \int_{-\infty}^{\infty} \Psi^* \rho \Psi \mathrm{d}x$$

which we denote by  $\langle \rho \rangle /m$ :

$$\langle p \rangle = m \frac{\mathrm{d} \langle x \rangle}{\mathrm{d}t} = \int_{-\infty}^{\infty} \Psi^* \rho \Psi \mathrm{d}x$$
 (2.17)

Thus, the expected or mean value of the momentum is related to the time-rateof-change of the mean position  $\langle x \rangle$  in a way consistent with (in correspondence with, as required) the classical notion of momentum.

The Heisenberg uncertainty principle can now be stated as,

$$\sigma_x \sigma_\rho \ge \frac{\hbar}{2} \tag{2.18}$$

where

$$\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2 , \qquad \sigma_\rho^2 = \langle p^2 \rangle - \langle p \rangle^2$$

## 2.9 Wave Packets and Fourier Transforms

Consider again the case of rectilinear motion of a free particle x in a vacuum. The wave function is

$$\Psi(x,t) = \psi(x)e^{i\omega t}$$
  
=  $\psi(x)e^{-iEt/\hbar}$  (2.19)

and  $\psi(x)$  satisfies

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = E\psi \tag{2.20}$$

The solution is

$$\Psi(x,t) = A \mathrm{e}^{\mathrm{i}k\left(x - \frac{\hbar kt}{2m}\right)} \tag{2.21}$$

where

$$k = \pm \sqrt{\frac{2mE}{\hbar}} \tag{2.22}$$

Clearly,  $E = \hbar k^2 / 2m = 2\pi^2 \hbar / m\lambda^2$ , so that the energy increases as the wave length  $\lambda$  decreases.

This particular solution is not normalizable  $(\int_{\mathbb{R}} \Psi^* \Psi dx \to \infty)$ , so that a free particle cannot exist in a stationary state. The spectrum in this case is continuous, and the solution is of the form,

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(k) \mathrm{e}^{\mathrm{i}(kx - \hbar k^2 t/2m)} \mathrm{d}k$$
(2.23)

$$\Psi(x,0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(k) \mathrm{e}^{\mathrm{i}kx} \mathrm{d}x \qquad (2.24)$$

$$\varphi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi(x,0) \mathrm{e}^{\mathrm{i}kx} \mathrm{d}x \qquad (2.25)$$

Equation (2.23) characterizes the wave as a sum of *wave packets*. It is a sinusoidal function modulated by the function  $\varphi$ . The wave function is the Fourier transform of  $\varphi$  and  $\varphi$  is the inverse Fourier transform of the initial value  $\Psi(x, 0)$  of  $\Psi$ . Instead of a particle velocity as in classical mechanics, we have a *group velocity* of the envelope of wavelets;

#### 2.10 The Wave-Momentum Duality

In classical mechanics, the dynamical state of a particle is defined at every instant of time by specifying its position  $\boldsymbol{x}$  and its momentum  $\boldsymbol{p}$ . In quantum mechanics, one can only define the probability of finding the particle in a given region when one carries out a measurement of the position. Similarly, one cannot define the momentum of a particle; one can only hope to define the probability of

#### 2.11. APPENDIX

determining the momentum in a given region of *momentum space* when carrying out a measure of momentum.

To define the probability of finding momentum  $\boldsymbol{p}$  in a volume  $(\boldsymbol{p}, \boldsymbol{d} + d\boldsymbol{p})$ , we consider, for fixed time t, the Fourier transform  $\Phi(\boldsymbol{p})$  of the wave function  $\Psi$ :

$$\Phi(\boldsymbol{p}) = \frac{1}{(2\pi\hbar)^{n/2}} \int_{\mathbb{R}^n} \Psi(\boldsymbol{x}) e^{-i\boldsymbol{p}\cdot\boldsymbol{x}} dx$$

$$\Psi(\boldsymbol{x}) = \frac{1}{(2\pi\hbar)^{n/2}} \int_{\mathbb{R}^n} \Phi(\boldsymbol{p}) e^{i\boldsymbol{p}\cdot\boldsymbol{x}} dp$$

$$(2.26)$$

(with  $dx = dx_1 \cdots dx_N$ ,  $dp = dp_1 \cdots dp_N$ ). Thus, the wave function can be viewed as a linear combination of waves  $\exp(i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar)$  of momentum  $\boldsymbol{p}$  with coefficients  $(2\pi\hbar)^{-n/2}\Phi(\boldsymbol{p})$ . The probability of finding a momentum  $\boldsymbol{p}$  in the volume  $(\boldsymbol{p}, d\boldsymbol{p})$ is

$$\pi(\boldsymbol{p}) = \Phi^*(\boldsymbol{p})\Phi(\boldsymbol{p})$$

and we must have

$$\int_{\mathbb{R}^n} \Phi^*(\boldsymbol{p}) \Phi(\boldsymbol{p}) \mathrm{d}\boldsymbol{p} = 1$$
(2.27)

Thus, the Fourier transform  $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$  establishes a one-to-one correspondence between the wave function and the momentum wave function. Equation (2.27) follows from Plancherel's identity,

$$\langle \Psi | \Psi \rangle = \| \Psi \|^2 = \langle \mathcal{F}(\Psi) | \mathcal{F}(\Psi) \rangle = \langle \Phi | \Phi \rangle$$
(2.28)

The interpretation of the probability densities associated with  $\Phi$  and  $\Psi$  is important. When carrying out a measurement on either position or momentum, neither can be determined with precision. The predictions of the probabilities of position and momentum are understood to mean that a very large number Nof equivalent systems with the same wave function  $\Psi$  are considered. A position measurement on each of them gives the probability density  $\Psi^*(\boldsymbol{x})\Psi(\boldsymbol{x})$  of results in the limit as N approaches infinity. Similarly,  $\Phi^*(\boldsymbol{x})\Phi(\boldsymbol{x})$  gives the probability density of results of measuring the momentum.

#### 2.11 Appendix

*Probability.* (Inspired by Griffith's Example [6, page 5])

Of 14 people in a room, the distribution of ages is as follows:



Figure 2.2: Histogram showing the number of people versus age.

- 1 aged 14 $\mathbf{2}$ aged 151 aged 161 aged 18 4aged 193 aged 231 aged 24
- The total number of people in the room

$$N = \sum_{j=0}^{\infty} N(j) = 13$$

• The Probability a person selected randomly is of age j

$$P(j) = \frac{N(j)}{N}$$

• The probability of getting either 14 or 15 is P(14) + P(15) = 3/13. The sum of all probabilities,

$$\sum_{j=0}^{\infty} P(j) = \frac{1}{N} \sum_{j=0}^{\infty} N(j) = 1$$
#### 2.11. APPENDIX

• The mean age, denoted  $\langle j \rangle$ :

$$\langle j \rangle = \frac{\sum j N(j)}{N} = \sum_{j=0}^{\infty} j P(j)$$

Check it out:

$$\frac{(14) + 2(15) + (16) + (18) + 4(19) + 3(23) + (24)}{13} = \frac{247}{13} = 19$$

$$= 14\left(\frac{1}{13}\right) + 15\left(\frac{2}{13}\right) + 16\left(\frac{1}{13}\right) + 18\left(\frac{1}{13}\right) + 19\left(\frac{4}{13}\right)$$

$$+ 23\left(\frac{3}{13}\right) + 24\left(\frac{1}{13}\right)$$

$$= 14P(14) + 15P(15) + 16P(16) + 18P(18) + 19P(19)$$

$$+ 23P(23) + 24P(24)$$

$$= \sum_{j=0}^{\infty} jP(j)$$

We easily verify that

$$\left\langle j^2 \right\rangle = \sum j^2 P(j)$$

In general

$$\langle f(j) \rangle = \sum_{j=0}^{\infty} f(j) P(j) =$$
 expected value of  $f$ 

The mean is the same (as is the most probable value). To distinguish between them, a measure of spread is needed.

• The spread is measured by the variance

$$\sigma^{2} = \left\langle (\Delta j)^{2} \right\rangle = \sum_{j=0}^{n} (j - \langle j \rangle)^{2} P(j)$$

and

$$\sigma = \sqrt{\langle (\Delta j)^2 \rangle}$$
 = the standard deviation.



**Figure 2.3**: Two histograms with the same median, average, and probable value, but different standard deviations.

• It is easy to show that

$$\sigma^2 = \left\langle j^2 \right\rangle - \left\langle j \right\rangle^2$$

For the continuous case,

$$P_{ab} = \int_{a}^{b} \rho(x) dx = \text{probability that } x \in [a, b]$$
$$\int_{-\infty}^{\infty} \rho(x) dx = 1$$
$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x) \rho(x) dx$$
$$\sigma^{2} = \langle x^{2} \rangle - \langle x \rangle^{2} , \qquad \langle x \rangle = \int_{-\infty}^{\infty} x \rho(x) dx$$

# Dynamical Variables and Observables in Quantum Mechanics: The Mathematical Formalism

### **3.1** Introductory Remarks

We have seen that the solutions of Schrödinger's equations are wave functions  $\Psi = \Psi(x,t)$  that have the property that  $|\Psi|^2 = \Psi^*\Psi(x,t)$  is the probability distribution function giving the probability that the elementary particle under study is at position x at time t (actually that the particle is in the volume between x and x + dx). Moreover, the knowledge of the wave function  $\Psi$  (or equivalently, the momentum wave function  $\Phi = \Phi(p,t)$ ) determines completely the dynamical state of the quantum system. We shall now build on these ideas and develop the appropriate mathematical setting for an operator theoretic framework for quantum mechanics that brings classical tools and concepts to the theory.

Everything we derive is applicable to functions defined on  $\mathbb{R}^N$ , but virtually all of the results can be demonstrated without loss in generality in  $\mathbb{R}$ . The notation for the spatial coordinate  $\mathbf{x}$  will be used interchangeably with  $\mathbf{q}$  (or xwith q in  $\mathbb{R}$ ), as  $\mathbf{q}$  is the classical notation for a generalized coordinate in "phase space" where coordinate - momentum pairs  $(\mathbf{q},\mathbf{p}) = (\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_N,\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_N)$ define dynamical states. We will frequently treat dynamical variables, such as momentum p, as operators, and when it is important to emphasize the operator character of the result, we will affix a tilde to the symbol (i.e. for momentum p, the associated operator is  $\tilde{p} = -i\hbar\frac{d}{dq}$ ). Thus, a function  $F = F(\mathbf{q},\mathbf{p})$  is associated with an operator  $\tilde{F}(\mathbf{q}_1,\mathbf{q}_2,...,\mathbf{q}_N,-i\hbar\frac{\partial}{\partial\mathbf{q}_1},-i\hbar\frac{\partial}{\partial\mathbf{q}_2},...,-i\hbar\frac{\partial}{\partial\mathbf{q}_n})$  etc. The coordinates  $(q_1, q_2, ..., q_N)$  will hereafter be understood to be cartesian coordinates because the operator notation must represent dynamical quantities in a way that is invariant under a change (e.g. a rotation) of the coordinate axes. Indeed, while ordinary multiplication of functions is commutative, the corresponding operators may not commute, so the  $q_i$  are interpreted as cartesian coordinates to avoid ambiguity.

# **3.2** The Hilbert Space $L^2(\mathbb{R})$ (or $L^2(\mathbb{R}^N)$ )

Since the wave function must have the property that  $|\Psi|^2$  is integrable over  $\mathbb{R}$ , ( $\Psi$  is "square-integrable"), it must belong to the following space:

 $L^2(\mathbb{R})$  (or  $L^2(\mathbb{R}^N)$ ) is the space of equivalence classes [u] of measurable complex-valued functions equal almost everywhere on  $\mathbb{R}$  (or  $\mathbb{R}^N$ )( $v \in [u] \Rightarrow v = u$  everywhere except a set of measure zero) such that  $|u|^2 = u^*u$  is Lebesgue integrable on  $\mathbb{R}$  (or  $\mathbb{R}^N$ ).

Thus,  $u \in L^2(\mathbb{R})$  implies that u represents an equivalence class [u] of functions equal almost everywhere on  $\mathbb{R}$  such that  $\int_{\mathbb{R}} u^2 dx < \infty$ .

The space  $L^2(\mathbb{R})$  is a complete inner product space with inner product

$$\langle \psi, \varphi \rangle = \int_{IR} \psi^* \varphi dx$$

where  $\psi^*$  is the complex conjugate of  $\psi$ . Thus  $L^2(\mathbb{R})$  is a Hilbert space. The associated norm on  $L^2(\mathbb{R})$  is then

$$\parallel \psi \parallel = \sqrt{\langle \psi, \psi \rangle}$$

It can be shown that  $L^2(\mathbb{R})$  is reflexive (in particular, by the Riesz theorem, for every continuous linear functional f in the dual  $(L^2(\mathbb{R}))'$ , there is a unique  $u_f \in L^2(\mathbb{R})$  such that  $f(v) = \langle v, u_f \rangle$  and  $|| f ||_{(L^2(\mathbb{R}))'} = || u_f ||)$ . Also,  $L^2(\mathbb{R})$ is separable, meaning that it contains countable everywhere dense sets. In other words, for any  $u \in L^2(\mathbb{R})$ , and any  $\varepsilon > 0$ , there exists an infinite sequence of functions  $\{u_n\}_{n=1}^{\infty}$  in  $L^2(\mathbb{R})$  and an integer M > 0 such that  $|| u_n - u || < \varepsilon$  for all n > M. We shall demonstrate how such countable sets can be computed given any u in the next section.

## 3.3 Dynamical Variables and Hermitian Operators

A dynamical variable is some physical feature of the quantum system that depends upon the physical state of the system. In general, we adopt the convention that the state is described by the position coordinates  $\mathbf{q}$  and the momentum  $\mathbf{p}$  (or  $\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N; \mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N$ ). Thus, a dynamical variable is a function

$$Q = Q(\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N; \mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N)$$

Since the momentum components  $\mathbf{p}_j$  can be associated with the operators  $\mathbf{p}_j = -i\hbar\partial/\partial \mathbf{q}_j$ , dynamical variables likewise characterize operators,

$$Q(\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N; \mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N) = \widetilde{Q}(\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N); -i\hbar \frac{\partial}{\partial \mathbf{q}_1}, -i\hbar \frac{\partial}{\partial \mathbf{q}_2}, ..., -i\hbar \frac{\partial}{\partial \mathbf{q}_N}, -i\hbar \frac{\partial}{\partial \mathbf{q}_N}, ..., -i\hbar \frac{\partial$$

(Analogously, one could define an operator

$$Q(\mathbf{q};\mathbf{p}) = \hat{Q}\left(i\hbar\frac{\partial}{\partial\mathbf{p}_1}, -i\hbar\frac{\partial}{\partial\mathbf{p}_2}, ..., -i\hbar\frac{\partial}{\partial\mathbf{p}_N}\right); \mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N$$

The *expected value* or *mean* of a dynamical variable Q for quantum state  $\Psi$  is denoted  $\langle Q \rangle$  and is defined as  $(\mathbb{R}^N, N = 1)$ ,

$$\langle Q \rangle = \int_{\mathbb{R}} \psi^* Q \psi dq / \int_{\mathbb{R}} \psi^* \psi dq$$
  
=  $\frac{\left\langle \Psi, \tilde{Q}\Psi \right\rangle}{\left\langle \Psi, \Psi \right\rangle}$  (3.1)

Any operator  $A: L^2(\mathbb{R}) \to L^2(\mathbb{R})$  is said to be *Hermitian* if

$$\langle \psi, A\varphi \rangle = \langle A\psi, \varphi \rangle^* \quad \forall \psi, \varphi \in L^2(\mathbb{R})$$
 (3.2)

Any operator Q that corresponds to a genuine dynamical variable of a physical system, must be such that its expected value  $\langle Q \rangle$  is real; i.e.,

$$\langle Q \rangle = \left\langle \psi, \hat{Q}\psi \right\rangle / \left\langle \psi, \psi \right\rangle$$

is real. Therefore, for such Q we must have

$$\left\langle \psi, \tilde{Q}\varphi \right\rangle = \left\langle \tilde{Q}\psi, \varphi \right\rangle \quad \forall \psi, \varphi \in L^2(\mathbb{R})$$

In other words, every dynamical variable must be characterized by a Hermitian operator on  $L^2(\mathbb{R})$  (or  $L^2(\mathbb{R}^N)$ ).

We note that the *variance* of a dynamical variable is defined as

$$\sigma_Q^2 = \left\langle Q^2 \right\rangle - \left\langle Q \right\rangle^2 \tag{3.3}$$

The dynamical variable Q takes on the numerical value  $\langle Q \rangle$  with certainty if and only if  $\sigma_Q^2 = 0$ .

This observation leads us to a remarkable property of those dynamical variables that can actually be measured in quantum systems with absolute certainty. Since

$$\begin{aligned} \sigma_Q^2 &= \langle Q^2 \rangle - \langle Q \rangle^2 \\ &= \frac{\langle \psi, Q^2 \psi \rangle}{\langle \psi, \psi \rangle} - \left( \frac{\langle \psi, Q \psi \rangle}{\langle \psi, \psi \rangle} \right)^2 \\ &= \frac{\langle Q \psi, Q \psi \rangle}{\langle \psi, \psi \rangle} - \frac{\langle \psi, Q \psi \rangle^2}{\langle \psi, \psi \rangle^2} \end{aligned}$$

If  $\sigma_Q^2 = 0$ , we have

$$\left\langle \psi, \tilde{Q}\psi \right\rangle^2 = \parallel \tilde{Q}\psi \parallel^2 \parallel \psi \parallel^2 \quad \forall \psi \in L^2(\mathbb{R})$$
 (3.4)

This is recognized as the Cauchy-Sworz inequality for  $\langle \psi, \tilde{Q}\psi \rangle$  in the case in which the equality holds, which happens when the functions  $Q\psi$  and  $\psi$  are proportional, i.e. when there exists a constant  $\lambda \in \mathbb{C}$  such that

$$\tilde{Q}\psi = \lambda\psi \tag{3.5}$$

This is recognized as the eigenvalue problem for the Hermitian operator  $\hat{Q}$ , with eigenvalue - eigenfunction pair  $(\lambda, \psi)$ . Therefore, the "fluctuations" (variances) of the dynamical variable Q from its mean  $\langle Q \rangle$  vanish for states  $\psi = \psi_{\lambda} \in L^2(\mathbb{R})$  which are eigenfunctions of the operator  $\tilde{Q}$ . Notice that if  $|\psi|^2 = 1$ , then  $\langle \psi, \tilde{Q}\psi \rangle = \langle Q \rangle = \lambda \langle \psi, \psi \rangle = \lambda$ ; i.e. the expected value  $\langle Q \rangle$  is an eigenvalue of  $\tilde{Q}$ . We summarize this finding in the following theorem,

**Theorem 1** The dynamical variable Q of a quantum system possesses with certainty (with probability 1) the well-defined value  $\langle Q \rangle$  if and only if the dynamical state of the system is represented by an eigenfunction  $\psi$  of the Hermitian operator  $\tilde{Q} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$  associated with Q; moreover, the expected value  $\langle Q \rangle$  is an eigenvalue of  $\tilde{Q}$ . This fundamental result establishes the connection of quantum mechanics with the spectral theory of Hermitian operators. We explore this theory in more detail for the case of a discrete spectrum.

## 3.4 Spectral Theory of Hermitian Operators: The Discrete Spectrum

Returning to the eigenvalue problem (3.5), let us consider the case in which there exists a countable but infinite set of eigenvalues  $\lambda_k$  and eigenfunctions  $\varphi_k \in L^2(\mathbb{R}), \ k = 1, 2, ...$  for the operator  $\tilde{Q}$ . In this case,  $\tilde{Q}$  is said to have a *discrete spectrum*. The basic properties of the system in this case are covered or derived from the following theorem,

**Theorem 2** Let  $(\lambda_k, \varphi_k), k = 1, 2, ...$  denote a countable sequence of eigenvalue - eigenfunction pairs for the Hermitian operator  $\tilde{Q} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ ; i.e.

$$\hat{Q}\varphi_k = \lambda_k \varphi_k, \quad k = 1, 2, \dots$$
 (3.6)

Then

- 1. if  $\varphi_k$  is an eigenfunction, so also is  $c \varphi_k$ , c being any constant;
- 2. if  $\varphi_k^{(1)}, \varphi_k^{(2)}, ..., \varphi_k^{(M)}$  are *M* eigenfunctions corresponding to the same eigenvalue  $\lambda_k$ , then any linear combination of these eigenfunctions is an eigenfunction corresponding to  $\lambda_k$ ;
- 3. the eigenvalues are real;
- 4. the eigenfunctions  $\varphi_k$  can be used to construct an orthonormal set, i.e., a set of eigenfunctions of  $\tilde{Q}$  such that

$$\langle \varphi_k, \varphi_m \rangle = \delta_{km} = \begin{cases} 1 & \text{if } k \neq m \\ 0 & \text{if } k = m \end{cases}$$
(3.7)

5. any state  $\psi \in L^2(\mathbb{R})$  can be represented as a series,

$$\psi(q) = \sum_{k=1}^{\infty} c_k \varphi_k(q) \tag{3.8}$$

where

$$c_k = \langle \psi, \varphi_k \rangle \tag{3.9}$$

and by (3.8) we mean

$$\lim_{m \to \infty} \| \psi - \sum_{k=1}^{m} c_k \varphi_k \| = 0$$
(3.10)

*Proof.* Parts 1 and 2 are trivial. In property 2, the eigenfunctions are said to have a *degeneracy* of order M.

To show 3, we take the inner product of both sides of (3.6) by  $\varphi_m$  and obtain the number

$$\lambda_{k} = \left\langle \varphi_{m}, \tilde{Q}\varphi_{k} \right\rangle / \left\langle \varphi_{m}, \varphi_{k} \right\rangle$$

which is real, because  $\tilde{Q}$  is Hermitian.

4. Consider,

$$Q \varphi_k = \lambda_k \psi_k$$
 and  $\psi_m \varphi m$ ,  $m \neq k$ 

Then,

$$\left\langle \varphi_m, \tilde{Q}\varphi_k \right\rangle = \lambda_k \left\langle \varphi_m, \varphi_k \right\rangle = \left\langle \tilde{Q}\varphi_m, \varphi_k \right\rangle = \lambda_m \left\langle \varphi_m, \varphi_k \right\rangle$$

Thus,  $(\lambda_k - \lambda_m) \langle \varphi_m, \varphi_k \rangle = 0$  for  $m \neq k, \ \lambda_k \neq \lambda_m$ .

5. The eigenfunctions  $\varphi_k$  are assumed to be normalized:

$$\langle \varphi_k, \varphi_k \rangle = 1$$
 ,  $k = 1, 2, \dots$ 

Thus they form an orthonormal basis for  $L^2(\mathbb{R})$ . Equation (3.8) is then just the Fourier representation of  $\psi$  with respect to this bases. Equation (3.9) follows by simply computing  $\langle \Sigma_m c_m \varphi_m, \varphi_k \rangle$  and using (3.7).

Various functions of the operator  $\tilde{Q}$  can likewise be given a spectral representation. For instance, if

$$\begin{split} \tilde{Q}\varphi_k &= \lambda_k \varphi_k, \text{ then } \tilde{Q}^2 \varphi_k &= \tilde{Q}(\tilde{Q}\varphi_k) \\ &= \tilde{Q}\lambda_k \varphi_k \\ &= \lambda_k \tilde{Q}\varphi_k \\ &= \lambda_k^2 \varphi_k, \end{split}$$

and, in general

$$(\tilde{Q})^r \varphi_k = \lambda_k^r \varphi_k$$

and, symbolically,

$$(\sin \tilde{Q})\varphi_k = (\tilde{Q} - \frac{1}{3!}\tilde{Q}^3 + \cdots)\varphi_k$$
$$= (\lambda_k - \frac{1}{3!}\lambda_k^3 + \cdots)$$
$$= \sin(\lambda_k)\varphi_k$$

etc. In general, if  $\tilde{Q}$  is a Hermitian operator with discrete eigenvalue - eigenfunction pairs  $(\lambda_k, \varphi_k)$  and F is any smooth function on  $\mathbb{R}$ , we may write

$$F(\tilde{Q}) = \sum_{k=1}^{\infty} F(\lambda_k) \varphi_k$$
(3.11)

For  $\tilde{Q}: L^2(\mathbb{R}) \to L^2(\mathbb{R})$ , (11) is meaningful if the series converges; i.e. if the sequence of real numbers,

$$\mu_n = \sum_{k=1}^n |F(\lambda_k)|^2 \text{ converges in } \mathbb{R}.$$

For

$$F(\tilde{Q}) = e^{i\xi\tilde{Q}}, \xi \in \mathbb{R},$$

this series always converges. In particular,

$$e^{i\xi\tilde{Q}}\psi = \sum_{k} c_{k}e^{i\xi\tilde{Q}}\varphi_{k}$$
$$= \sum_{k} \sum_{m} c_{k}e^{i\xi\lambda m}\varphi_{m}\varphi_{k}$$

and  $|c_k e^{i\xi\lambda m}|^2 = |c_k|^2 \to 0$  as  $m \to \infty$ .

### **3.5** Observables and Statistical Distributions

The statistical distribution of a quantity Q associated with a quantum dynamical system is established by its *characteristic function*  $a : \mathbb{R} \to \mathbb{R}$ . In general, the characteristic function is, to within a constant, the Fourier transform of the probability density function  $\rho$  associated with a random variable X:

$$a(\xi) = \int_{-\infty}^{\infty} e^{i\xi x} \varrho(x) dx \tag{3.12}$$

 $(\varrho(x))$  being the probability of finding X in (x, x+dx). Thus,  $a(\xi)$  is the expected value of  $e^{i\xi x}$ :

$$a(\xi) = \left\langle e^{i\xi x} \right\rangle \tag{3.13}$$

If the random variable X can only assume discrete values  $x_k$ ,  $k = 1, 2, \cdots$ , and if  $\rho_1, \rho_1, \cdots$  are the probabilities of these values, then

$$a(\xi) = \sum_{k=1}^{\infty} \varrho_k e^{i\xi x_k} \tag{3.14}$$

with  $\sum_{k} \varrho_k = 1$ .

Returning now to quantum dynamics, we shall refer to any dynamical variable Q of a quantum system characterized by a Hermitian operator  $\tilde{Q}$  on  $L^2(\mathbb{R})$  with a discrete spectrum of eigenvalues  $\lambda_k$  as an *observable*. The reason that term is appropriate becomes clear when we consider the characteristic function of the quantum system. For the wave function  $\Psi$  representing the dynamical state of the system,

$$a(\xi) = \frac{\left\langle \Psi, e^{i\xi\tilde{Q}}\Psi \right\rangle}{\left\langle \Psi, \Psi \right\rangle}$$
$$= \sum_{k=1}^{\infty} \varrho_k e^{i\xi\lambda_k}$$
(3.15)

where now

$$\varrho_{k} = |c_{k}|^{2} / \langle \Psi, \Psi \rangle 
= |\langle \varphi_{k}, \Psi \rangle|^{2} / \langle \Psi, \Psi \rangle$$
(3.16)

Comparing (3.14) and (3.15), we arrive at the following theorem:

#### 3.6. THE CONTINUOUS SPECTRUM

**Theorem 3** Let Q be an observable of a quantum dynamical system with the discrete spectrum of eigenvalues  $\{\lambda_k\}_{k=1}^{\infty}$ . Then

1. The only values Q can assume are precisely the eigenvalues

$$\lambda_1, \lambda_2, \cdots \lambda_m, \cdots, and$$

2. the probability that Q takes on the value  $\lambda_k$  is

$$\varrho_k = |c_k|^2 = |\langle \varphi_k, \Psi \rangle|^2$$

where  $\varphi_k$  is the eigenfunction corresponding to  $\lambda_k$  and  $\Psi$  is the normalized wave function,  $\langle \Psi, \Psi \rangle = 1$ .

We observe that the discrete probabilities  $\varrho_k$  satisfy

$$1 = \sum_{k} \varrho_{k} = \sum_{k} \frac{|\langle \varphi_{k}, \Psi \rangle|^{2}}{\langle \Psi, \Psi \rangle} = \sum_{k} |c_{k}|^{2}$$

as required.

### 3.6 The Continuous Spectrum

Not all dynamical variables (Hermitian operators) have a discrete spectrum; i.e. not all are observables in quantum dynamical systems. For example, the momentum operator  $p = -i\hbar d/dq$  is Hermitian but does not have a discrete spectrum and, therefore, is not an observable (in keeping with our earlier view of the uncertainty principle for position and momentum). In particular, the eigenvalue problem for p is (Cf. Messiah [13] p.178)

$$pU(p',q) = p'U(p',q)$$
(3.17)

where U(p',q) is the eigenfunctions of p with an eigenvalue  $\lambda = p'$ , a continuous function of the variable p'; i.e. the spectrum of  $p = \frac{\hbar}{i} \frac{d}{dq}$  is continuous. We easily verify that

$$U(p',q) = \frac{1}{\sqrt{2\pi\hbar}} e^{ip'q/\hbar}$$
(3.18)

By analogy with the Fourier series representation (3.8) of the dynamical state  $\psi$  for discrete spectra, we now have the Fourier transform representation of the state for the case of a continuous spectrum:

$$\psi(q) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \varphi(p') e^{ip'q/\hbar} dp'$$
(3.19)

In analogy with (3.9),

$$\varphi(p') = \left\langle U(p',q), \psi(q) \right\rangle \tag{3.20}$$

The analogue arguments cannot be carried further. Indeed, the function (3.18) is not in  $L^2(\mathbb{R})$  and a more general setting must be constructed to put the continuous spectrum case in the proper mathematical framework.

## 3.7 The Generalized Uncertainty Principle for Dynamical Variables

Recall that for any dynamical variable characterized by a Hermitian operator Q, its expected value in the state  $\Psi$  is

$$\langle Q \rangle = \left\langle \Psi, \tilde{Q}\Psi \right\rangle$$

where  $\tilde{Q}$  is the operator associated with  $Q : (Q(x, p, t) = \tilde{Q}(x, \frac{\hbar}{i} \frac{\partial}{\partial x}, t))$ , and  $\langle \Psi, \Psi \rangle = 1$ . Recall that its standard deviation (actually, its variance) is

$$\sigma_Q^2 = \left\langle (\tilde{Q} - \langle Q \rangle^2) \right\rangle = \left\langle \Psi, (\tilde{Q} - \langle Q \rangle^2) \Psi \right\rangle$$
(3.21)

For any other such operator M,

$$\sigma_M^2 = \left\langle \Psi, (\tilde{M} - \langle M \rangle^2) \Psi \right\rangle$$
$$= \left\langle (\tilde{M} - \langle M \rangle^2) \Psi, (\tilde{M} - \langle M \rangle^2) \Psi \right\rangle$$

Noting that for any complex number z,  $|z|^2 \ge (I_m(z))^2 \ge [\frac{1}{2i}(z-z)]^2$ , it is an algebraic exercise to show that

$$\sigma_Q^2 \sigma_M^2 \ge \left(\frac{1}{2i} \left\langle \tilde{Q}\tilde{M} - \tilde{M}\tilde{Q} \right\rangle \right)^2 \tag{3.22}$$

where [Q, M] is the commutator of the operators Q and M:

$$[Q, M] = \tilde{Q}\tilde{M} - \tilde{M}\tilde{Q} \tag{3.23}$$

The generalized Heisenberg uncertainty principle is this: for any pair of incompatible observables (those for which the operators do not commute, or  $[Q, M] \neq 0$ ),

#### 3.7. THE GENERALIZED UNCERTAINTY PRINCIPLE

the condition (3.22) holds. Such incompatible observables cannot share common eigenfunctions.

To demonstrate that (3.22) is consistent with the elementary form of the uncertainty principle discussed earlier, set  $\tilde{Q} = x$  and  $\tilde{M} = \tilde{p} = (\hbar/i)d/dx$ . Then, for a test state  $\psi$ ,

$$\begin{bmatrix} \tilde{x}, \tilde{p} \end{bmatrix} \psi = x \frac{\hbar}{i} \frac{d\psi}{dx} - \frac{\hbar}{i} \frac{d}{dx} (x\psi)$$
  
=  $i\hbar\psi$ 

Hence,

$$\sigma_x^2 \sigma_p^2 \ge \left(\frac{1}{2i}ih\right)^2 = \left(\frac{\hbar}{2}\right)^2$$

or

$$\sigma_x \sigma_p \ge \frac{\hbar}{2}$$

in agreement with the earlier calculation.

In the case of a Hermitian (or self-adjoint) operator Q with eigenvalue-eigenvector pairs  $\{(\lambda_k, \varphi_k)\}_{k=1}^{\infty}$ , we have seen that

$$Q\psi = \sum_{k=1}^{\infty} \lambda_k c_k \varphi_k, \quad c_k = \langle \psi, \varphi_k \rangle$$

The projection operators  $\{P_k\}$  defined by

$$P_k\psi = \langle \psi, \varphi_k \rangle \, \varphi_k = c_k \varphi_k$$

have the property that

$$\psi = I\psi = \sum_{k=1}^{\infty} c_k \varphi_k = \sum_{k=1}^{\infty} P_k \psi$$

so that, symbolically,

$$\sum_{k=1}^{\infty} P_k = 1 = \text{ the identity.}$$
(3.24)

For this reason, such a set of projections is called a *resolution of the identity*. The operators Q can thus be represented as

$$Q = \sum_{k=1}^{\infty} \lambda_k P_k$$

The same type of construction can be extended to cases in which Q has a continuous spectrum. In such cases, for any self-adjoint operator Q, there exists a unique family of projection  $\{P_{\lambda}\}, -\infty < \lambda < + < \infty$ , such that  $P_{\lambda} = 0$  for  $\lambda < m$ ,  $P_{\lambda} = I$  for  $\lambda > m$ , m and M being lower and upper bounds of Q, and

$$\lim_{\lambda \to M} \| P_{\lambda} \psi - P_M \psi \| = 0 \quad \forall \psi \in L^2(\mathbb{R})$$

or, in particular,  $\lim_{\lambda \to M^+} || P_{\lambda} \psi - 1 \psi || = 0$ . For this reason the family is still referred to as a resolution of the identity. In this case, we write

$$Q = \int_{m}^{M+\varepsilon} \lambda dP_{\lambda} \quad , 0 < \varepsilon < 1 \tag{3.25}$$

in analogy with (3.24).

Likewise, if f is a continuous complex-valued function defined on the interval  $[m, \hat{M} + \varepsilon]$ , we write

$$f(Q) = \int_{m}^{M+\varepsilon} f(\lambda)dP_{\lambda}$$
(3.26)

The integral in (3.26)(and (3.25)) is understood as the limit of the sum  $\sum_{k} f(\lambda_k)(P_{\lambda_k} - P_{\lambda_{k-1}}) \text{ on a partition } m = \lambda_0 < \lambda_1 < \dots < \lambda_n = M \text{ of } [m, M + \varepsilon]$ as  $\max_k (\lambda_k - \lambda_{k-1}) \to 0.$ 

Returning to the Hamiltonian operator H, if

$$\langle H\psi,\varphi\rangle = \int_{-\omega}^{\infty} \lambda d(P_{\lambda}\psi,\varphi)$$
 (3.27)

for all  $\psi, \varphi \in L^2(\mathbb{R})^{-\infty}$ , the symbolic relation,

$$H = \int_{-\omega}^{\infty} \lambda dP_{\lambda} \tag{3.28}$$

can be written in analogy to the spectral representation in the discrete spectrum case. For any such dynamical variable characterized by a Hermitian operator Q (or, equivalently,  $\tilde{Q}$ ), if  $dP_{\lambda}\psi \to 0$  as  $\lambda \to_{-}^{+} \infty$  or  $dP_{\lambda}\psi \to \psi$ , and if  $dP_{\lambda}\psi \to dP_{\lambda_{0}}\psi$  for  $\lambda \to \lambda_{0}$ ,  $\lambda \geq \lambda_{0}$ , (with  $dP_{\lambda_{1}}, \leq dP_{\lambda_{2}}$  for  $\lambda_{1} \leq \lambda_{2}$ ), we can write

$$\left\langle \tilde{Q}\psi,\varphi\right\rangle = \int_{-\omega}^{+\infty} \lambda(dP_{\lambda}\psi,\varphi)$$
 (3.29)

Chapter	4
---------	---

# Applications: The Harmonic Oscillator and the Hydrogen Atom

### 4.1 Introductory Remarks

In this chapter, we consider applications of the theory discussed thus far to basic examples in quantum mechanics. These include the elementary example of the case of the harmonic oscillator for which the notion of ground states and quantization of energy levels is immediately derivable from Schrödinger's equation. Then the calculation of the quantum dynamics (the wave function) for the fundamental problem of the hydrogen atom is also described.

## 4.2 Ground States and Energy Quanta: the Harmonic Oscillator

Let us return to the simple case of a particle moving along a straight line with position coordinate x = q. We may then consider wave functions of the form,

$$\Psi(q,t) = \psi(q)e^{iEt/\hbar} \tag{4.1}$$

Then Schrödinger's equation reduces to

$$\left(H\left(\frac{\hbar}{i}\frac{d}{dq},q\right) - E\right)\psi(q) = 0 \tag{4.2}$$

Thus, the energy E is an eigenvalue of the Hamiltonian H.

The Hamiltonian, we recall, is of the form

$$H(p,q) = \frac{p^2}{2m} + V(q) = -\frac{\hbar^2}{2m} \frac{d^2}{dq^2} + V(q)$$
(4.3)

Hence, to complete the definition of the operator  $\tilde{H}$ , we need to specify the potential V. A classical and revealing example concerns the case in which V represents the potential energy of a harmonic oscillator vibrating with angular frequency  $\omega$ :

$$V(q) = \frac{1}{2}m\omega^2 q^2 \tag{4.4}$$

Then (4.2) becomes

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dq^2} + \frac{1}{2}m\omega^2 q^2\right)\psi = E\psi \tag{4.5}$$

or

$$\left(\frac{d^2}{dq^2} + \lambda - \alpha^2 q^2\right)\psi = 0 \tag{4.6}$$

where

$$\lambda = 2mE/\hbar^2, \quad \alpha = m\omega/\hbar \tag{4.7}$$

The solutions of (4.6) are of the form

$$\psi(q) = ae^{-\alpha q^2/2}$$

and a direct substitution reveals that the corresponding eigenvalue is

$$\lambda = \alpha = mv/\hbar$$

Thus, the corresponding energy is

$$E = E_0 = \frac{1}{2}\hbar\omega \tag{4.8}$$

This lowest possible energy corresponds to the *ground state*, and is seen to be half of Planck's energy quantum.

The remaining solutions of the wave equation can be found by standard powerseries expansions (method of Frobenius) to be of the form

$$\psi_n(q) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \cdot \frac{1}{\sqrt{n!2^n}} H_n(\xi) e^{-\xi^2/2}$$
(4.9)

#### 4.3. THE HYDROGEN ATOM

where  $H_n(\cdot)$  are Hermite polynomials of order n and  $\xi = q(m\omega/\hbar)^{1/2}$ . The corresponding energy states are

$$E_n = (n + \frac{1}{2})\hbar\omega$$
,  $n = 0, 1, \cdots$  (4.10)

Thus, energy occurs in *quantized states* distinguished by different integer values of n, and the lowest energy is the ground state corresponding to n = 0. This elementary example shows how the realization of discrete quantum states is a natural mathematical consequence of the discrete spectrum of this second-order, linear, elliptic differential operator.

### 4.3 The Hydrogen Atom

The harmonic oscillator provides an elementary example of how ground and quantized energy states are natural mathematical properties of eigenvalues of Schrödinger's equation, but the particular choice of the potential V was, to an extent, contrived. We now turn to one of the most important examples in quantum mechanics that provides the building blocks for quantum chemistry: the Schrödinger equation for the hydrogen atom. This involves a Hamiltonian with a potential that characterizes the energy of a charged particle, an electron, orbiting a nucleus with a charge of equal magnitude but opposite sign.

We begin by writing the Schrödinger equation for the wave function in three dimensions for a single particle as follows:

$$i\hbar\frac{\partial\Psi}{\partial t} + \frac{\hbar^2}{2m}\Delta\Psi - V\Psi = 0 \tag{4.11}$$

where  $\triangle$  is the three-dimensional Laplacian,

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$
(4.12)

Setting

$$\Psi_n(x, y, z, t) = \psi_n(x, y, z)e^{-iE_n t/\hbar}$$
(4.13)

$$-\frac{\hbar^2}{2\hat{m}} \Delta \psi_n + V \psi_n = E_n \psi_n \tag{4.14}$$

where  $\hat{m} = m_e m_n / (m_e + m_n)$ ,  $m_e$  being the mass of the electron and  $m_n$  the mass of the nucleus. Since  $m_n = 1837m_e$ ,  $\hat{m} \approx 0.99946m_e$ , so hereafter we will write m for  $\hat{m}$  as a good approximation.

The general solution is, in principle, obtained by superposition:  $\Psi = \Sigma_n \Psi_n$ . But in this case, details of the properties of these eigenfunctions and eigenvalues play a fundamental role in shaping our understanding of the atomic structure of matter.

Because of spherical symmetry, it is convenient to transform the Schrödinger equation (4.14) into spherical coordinates  $(r, \theta, \phi)$ , r being the radial distance from the nucleus at the origin,  $\theta$  being the azimuthal angle with respect to the z-axis, and  $\phi$  the angle from the x-axis to the xy-plane. In these coordinates, the time-independent Schrödinger equation becomes

$$-\frac{\hbar^2}{2m} \left[ \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} \right] + V\psi = E\psi$$
(4.15)

For center-of-mass coordinates, with origin at the nucleus (r = 0), the potential V is

$$V(r) = -\frac{e^2}{4\pi_0\varepsilon_0 r} \tag{4.16}$$

We attempt to solve (4.15) by the method of separation of variables, wherein we separate the radial dependence from the  $\theta$ ,  $\phi$ -dependence by assuming that  $\psi$  is of the form

$$\psi(r,\theta,\phi) = R(r)Y(\theta,\phi) \tag{4.17}$$

The standard process is to substitute (4.17) into (4.15) and then divide by RY, and multiply by  $-2mr^2/\hbar^2$  to obtain the sum,

$$\mathcal{A}(R) + \mathcal{B}(Y) = 0 \tag{4.18}$$

where  $\mathcal{A}$  and  $\mathcal{B}$  are the differential expressions,

$$\mathcal{A}(R) = R^{-1} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) - \frac{2mr^2}{\hbar^2} [V(r) - E]$$

$$\mathcal{B}(Y) = Y^{-1} \left[ (\sin\theta)^{-1} \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial Y}{\partial \theta} \right) \right]$$
(4.19)

$$+ (\sin^2 \theta)^{-1} \frac{\partial^2 Y}{\partial \phi^2}$$

$$(4.20)$$

The standard argument now employed is that a function  $\mathcal{A}(R)$  of only r can equal a function  $-\mathcal{B}(Y)$  of only  $\theta$  and  $\phi$  only if they both are equal to a constant,

#### 4.3. THE HYDROGEN ATOM

say k, which we will show later must be an integer. We therefore arrive at two separate equations,

$$\mathcal{A}(R) = k$$
  
$$\mathcal{B}(Y) = -k \tag{4.21}$$

The Radial Equation. Since R(r) must be finite at r = 0 and must vanish as  $r \to \infty$ , we can choose the radial function to be of the form,  $R = a_0 e^{-\alpha r}$ ,  $a_0$  and  $\alpha$  being constants. Introducing this into (4.21) gives

$$\left(\alpha^2 + \frac{2mE}{\hbar^2} - \frac{2\alpha}{r} + \frac{2me^2}{4\pi\varepsilon_0\hbar^2r} - \frac{k}{r^2}\right)R(r) = 0$$

Upon equating like powers of r, we conclude that this equation is satisfied for  $R(r) \neq 0$  if k = 0 and

$$\alpha = \frac{me^2}{4\pi\varepsilon_0\hbar^2}$$

from which we also obtain

$$E = E_1 = -\frac{me^4}{(4\pi\varepsilon_0)^2 2\hbar^2}$$
(4.22)

The energy  $E_1$  in (4.22) is the ground state energy of the hydrogen atom.

Other radial solutions are obtained by considering trial solutions of the form  $(r+r^2+r^3+\cdots)e^{-\alpha r}$ . Replacing k by the integers  $k = \ell(\ell+1), \ \ell = 0, 1, \cdots, n-1$ , we can write the general solutions of  $\mathcal{A}(R) = \ell(\ell+1)$  as

$$R_{n\ell}(r) = -\left[\frac{4(n-\ell-1)!}{n^4 a_0^3[(n+\ell)!]^3}\right]^{1/2} \cdot \left(\frac{2r}{na_0}\right)^{\ell} \\ \times e^{-r/na_0} L_{1+n}^{1+2\ell} \left(\frac{2r}{na_0}\right) \\ n = 1, 2, \cdots, \ell = 0, 1, \cdots, n-1$$

$$(4.23)$$

where  $a_0$  is the so-called *Bohr radius* (because it is the radius of the orbit of lowest energy in Bohr's 1913 model of the hydrogen atom),

$$a_0 = \frac{4\pi\varepsilon_0\hbar^2}{me^2} \tag{4.24}$$

 $(a_0 \approx 0.529 \text{\AA})$ , and  $L_i^i(\cdot)$  are the Laguerre polynomials:

Examples are:

$$L_0^0 = 1, \ L_0^2 = 2, \ L_1^0 = 1 - x, \ L_1^2 = 18 - 6x, \ L_2^0 = x^2 - 4x + 2,$$

etc.

The indices n are called the *principal quantum numbers* and  $\ell$  are called the *azimuthal quantum numbers*. The energy eigenvalues are independent of  $\ell$  and are given by

$$E_{n} = -\frac{1}{n^{2}} \left[ \frac{me^{4}}{2\hbar^{2}(4\pi\varepsilon_{0})^{2}} \right]$$
  
=  $-\frac{E_{1}}{n^{2}}$ ,  $n = 1, 2, \cdots$  (4.26)

The formula (4.26) is called the *Bohr formula*. From it, we can calculate the wave length  $\lambda$  associated with quantum states n' and n''. According to Planck's formula, the energy of a photon is proportional to its frequency:  $E = hv = ch/\lambda$ . So

$$\frac{1}{\lambda} = R_0 \left( \left( \frac{1}{n''} \right)^2 - \left( \frac{1}{n'} \right)^2 \right)$$
(4.27)

where  $R_0 = (m/4\pi c\hbar^3) \cdot (e^2/4\pi\varepsilon_0)^2 = 1.097 \times 10^7 m^{-1}$ . The constant  $R_0$  is known as *Rydberg's constant* and (4.27) is *Rydberg's formula* for the hydrogen spectrum.

The so-called spectra (or line spectra) of atoms describes the electromagnetic radiation that atoms emit when excited. For the hydrogen atom, the spectrum characterizes transitions between its possible energy states. The energy differences between states is released as a photon of energy hv with wave number k = v/c. For the transition,  $n_1 \rightarrow n_2$ , the wave number of the emitted radiation is  $R_0((1/n_1)^2 - (1/n_2)^2)$ . For  $n_1$  fixed, and  $n_2 = 2, 3, \cdots$ , we obtain the so called Lyman series for ultraviolet light frequencies; for  $n_1 = 2, n_2 = 3, \cdots$ , the visible light, Balmer series is obtained. Infrared begins with the Brackett series at  $n_1 = 4, n_2 = 5, 6, \cdots$ . When the electron is removed, the ionized state is obtained. The minimum energy required to ionize the atom is *ionization energy*, which is  $hcR_0$ .

The Angular Equation. We return now to the angular equation,  $(4.21)_2$ , with  $k = \ell(1 + \ell)$ . Again we employ separation of variables, setting

$$Y(\theta\phi) = \Theta(\theta)\Phi(\phi)$$

We separate  $\mathcal{B}(Y) = -\ell(\ell+1)$  into differential equations in  $\theta$  and  $\phi$  with separation constant  $m^2$ , m found to be an integer by standard arguments: m =

#### 4.3. THE HYDROGEN ATOM

 $0, \pm 1, \pm 2, \cdots$ . The integers *m* are called *magnetic quantum numbers*. Solving the pair of separated equations, we find that

$$\Theta(\theta) = AP_{\ell}^{m}(\cos \theta) , \ \ell = 0, 1, \cdots, n-1$$
$$m = 0, \pm 1, \pm 2, \cdots$$
$$\Phi(\phi) = e^{im\phi}$$

where A is a constant and  $P_{\ell}^m$  are the Legendre polynomials:

$$P_{\ell}^{m}(x) = (1 - x^{2})^{|m|/2} \left(\frac{d}{dx}\right)^{|m|} P_{\ell}(x)$$
$$P_{\ell}(x) = \frac{1}{2!\ell!} \left(\frac{d}{dx}\right)^{\ell} (x^{2} - 1)^{\ell}$$
$$x \in [-1, 1]$$

Thus, the angular solutions are,

$$Y_{m\ell}(\theta, \phi) = \Theta_{m\ell}(\theta) \Phi_m(\phi) = \delta_m \sqrt{\frac{(1+2\ell)(\ell-|m|)!}{4\pi(1+|m|)!}} e^{im\phi} P_\ell^m(\cos\theta) \delta_m = (-1)^m \text{ for } m \ge 0, \ \delta_m = 1 \text{ for } m \le 0 \ell = 0, 1, \cdots, n-1; \ m = 0, \pm 1, \pm 2, \cdots$$
(4.28)

The functions  $Y_{m\ell}$  are called *spherical harmonics*. They are mutually orthogonal in the sense that

$$\int_0^{2\pi} \int_0^{\pi} Y_{m\ell}(\theta,\phi)^* Y_{m'\ell'}(\theta,\phi) \sin \theta d\theta d\phi$$
$$= \delta_{mm'} \delta_{\ell\ell'}$$

The Orbitals of the Hydrogen Atom. Summing up, the complete hydrogen wave functions, called *orbitals*, are

$$\begin{aligned}
\psi_{n\ell m} &= R_{n\ell}(r) Y_{\ell m}(\theta, \psi) \\
n &= 1, 2, \cdots \\
\ell &= 0, 1, \cdots, n-1 \\
m &= 0, \pm 1, \pm 2, \cdots
\end{aligned}$$
(4.29)

Quantum	Number	Orbital Name	Degeneracy	Energy
n	$\ell$			
1	0	1s	1	$E_1$ ((4.22))
2	0	2s	1	$E_{1}/4$
2	1	2p	3	$E_{1}/4$
3	0	3s	1	$E_{1}/9$
3	1	3p	3	$E_{1}/9$
÷		•	•	•

where  $R_{n\ell}(r)$  is given by (4.23) and  $Y(\theta, \phi)$  by (4.28). Again, these orbitals are mutually orthonormal:

$$\int_0^\infty \int_0^{2\pi} \int_0^\pi \psi_{n'\ell'm'}^* \psi_{n\ell m} r^2 dr \sin\theta d\theta d\phi$$
$$= \delta_{n'n} \delta \ell' \ell \delta_{m'm}$$

Spectroscopic States. The orbitals  $\psi_{n\ell m}$  are eigenvalues of the Hamiltonian for the hydrogen atom (assuming  $\hat{m} \approx m$ ) and the eigenvalues correspond to energies of the orbital states. Eigenfunctions with degenerate eigenvalues correspond to cases in which the values of two or more eigenvalues coincide for the same eigenfunction. As a means for categorizing various orbitals and energy states and their energy levels, a universally-used spectroscopic state convention is employed which assigns orbital types to various quantum numbers. The following conventions are used [15, page 94]:

The ground state energy  $E_1$  is negative, so the energies in the above table increase with increasing principal quantum number n.

Some important observations are listed as

1. The radial probability  $p_n \ell(r)$  is given by

$$p_{n\ell} = \int_0^{2\pi} \int_0^{\pi} r^2 \int_0^{\pi} r^2 |\psi_{n\ell m}|^2 \sin \theta d\theta d\phi$$
  
=  $r^2 R_{n\ell}^2(r)$ 

which decays rapidly as r increases. The portion of the domain in  $\mathbb{R}^3$  of each orbital containing, say, 90% of the probability density inside the boundary, define standard geometric shapes of the orbitals for various energy levels.

#### 4.3. THE HYDROGEN ATOM

- 2. For  $\ell = 0$ , we obtain the s-orbitals,  $1s, 2s, \dots, ns$  for various values of the principal quantum number n. These are spherical domains (under the conditions of the convention 1 above), with 1s corresponding to the (lowest) ground state energy. Only the s-orbitals are nonzero at r = 0.
- 3. Orbitals for the same value of n but different  $\ell$  and m are degenerate. For  $\ell \neq 0, m \neq 0$ , the so-called z-orbitals are obtained (denoted  $2p_z, 3d_z, \cdots$ ). The 2p orbitals, for example, comprise one real and two complex-valued functions,  $2p(+1) = \sqrt{3/4\pi}R(r)\sin\theta e^{i\phi}, 2p(0) = \sqrt{3/4\pi}R(r)\cos\theta$ , and  $2p(-1) = \sqrt{3/4\pi}R(r)\sin\theta e^{-i\phi}$ . The 2p(0) corresponds to the 2p orbital, and linear combinations, of  $2p(+1)^z$  and 2p(-1) correspond to the  $2p_x$  and  $2p_y$  orbitals.

$$2p_x = \frac{1}{2}[2p(+1) + 2p(-1)]$$
  
$$2p_y = \frac{1}{2}[2p(+1) + 2p(-1)]$$

Similar linear combinations of other degenerate pair lead to the  $3p, 3d, 4f, \cdots$  orbitals.

Examples of the hydrogen orbitals are illustrated in Fig.4.1

CHAPTER 3. THE MATHEMATICAL FORMALISM

Chapter 5

# Spin and the Pauli Principle

### 5.1 Angular Momentum and Spin

Up to this point, the *linear* momentum  $\mathbf{p}$  of a system has played the role of a fundamental dynamical property of quantum systems. From classical mechanics, however, we know that other types of momenta can exist by virtue of the moment of momentum vectors relative to a fixed point in space and by virtue of their spin about a trajectory of the particle. We shall now examine how these properties manifest themselves in quantum systems.

Consider an elementary particle located at position  $\mathbf{q}$  relative to a fixed origin and endowed with a momentum  $\mathbf{p}$ . The angular momentum of the particle consists of an extrinsic momentum  $\mathbf{L}$  due to its orbital motion about the origin, and an intrinsic momentum  $\mathbf{S}$  due to its spin about its center of mass. In classical mechanics,

$$\mathbf{L} = \mathbf{q} \times \mathbf{p} \quad \text{and} \; \mathbf{S} = I\boldsymbol{\omega}$$
 (5.1)

where I is an inertial quantity (such as the moment of inertia) and  $\boldsymbol{\omega}$  is the spin velocity vector.

Let us begin with the angular momentum  $\mathbf{L}$  of a particle (e.g. an electron of a hydrogen atom) and suppose that the linear momentum of the particle is  $\mathbf{p}$ . The corresponding momentum operator is

$$\tilde{p}_j = \frac{\hbar}{i} \frac{\partial}{\partial q_j} , \ j = 1, 2, 3$$
(5.2)

and the components of the angular momentum  $\mathbf{L} = \mathbf{q} \times \mathbf{p}$ , which we call the *extrinsic* angular momentum, are

$$L_j = \varepsilon_{rsj} q_r \frac{i}{\hbar} \frac{\partial}{\partial q_s} \tag{5.3}$$

# CHAPTER 5. FROM APPROXIMATE THEORIES TO MOLECULAR DYNAMICS

where repeated indices are summed and  $\varepsilon_{rsj}$  is the permutation symbol ( $\varepsilon_{rsj} = +1$  if rsj is an even permutation of integers  $1, 2, 3, \varepsilon_{rsj} = -1$  if rsj is an odd permutation, and  $\varepsilon_{rsj} = 0$  if rsj is not a permutation of 1, 2, 3).

The properties of the extrinsic angular momentum operator  $\mathbf{L} = \mathbf{q} \times \mathbf{p}$  are important because analogous properties hold for the intrinsic angular momentum, the spin  $\mathbf{S}$ , to be introduced momentarily.

From (5.3), we easily show that

$$[L_1, L_2] = i\hbar L_3 , [L_2, L_3] = i\hbar L_1 , [L_3, L_1] = i\hbar L_2$$
(5.4)

The operators  $L_i$  are therefore incompatible observables, and

$$\sigma_{L_1}^2 \sigma_{L_2}^2 \ge \left(\frac{1}{2i} \left\langle i\hbar L_3 \right\rangle\right)^2 = \frac{\hbar^2}{4} \left\langle L_3 \right\rangle^2 \quad , \text{ etc.}$$
(5.5)

On the other hand, the square of the angular momentum,

$$L^2 = L_i L_i = L_1^2 + L_2^2 + L_3^2 (5.6)$$

does commute and

$$[L^2, L_k] = 0$$
,  $k = 1, 2, 3$  (5.7)

Therefore, we can hope to find simultaneous eigenstates of  $L^2$  and  $L_i$ ; i.e. states  $\varphi$  such that

$$L^2 \varphi = \lambda \varphi$$
 and  $L_i \varphi = \mu \varphi$  (5.8)

It is perhaps not surprising that the eigenfunctions  $\varphi$  of the extrinsic angular momentum operators  $L^2$  and  $L_i$  for a single particle are none other than the angular and azimuthal parts of the orbitals for the single-electron hydrogen atom discussed earlier; i.e., the spherical harmonics  $Y_{\ell}(\theta, \phi)$  given by (4.27) for specific ranges of  $\ell$  and m. In particular,

$$L^{2}Y_{\ell m} = \hbar^{2}\ell(1+\ell)Y_{\ell m} \text{ and} 
L_{i}Y_{\ell m} = \hbar m Y_{\ell m} \text{ for} 
\ell = 0, 1/2, 3/2, \cdots 
m = -\ell, -\ell + 1, \cdots, \ell - 1, \ell 
i = 1, 2, 3$$
(5.9)

Spin. In addition to its extrinsic angular momentum owning to its angular motion about the nucleus, described above, an electron also is endowed with an intrinsic spin momentum analogous to the spin of a body in motion about an axis along its trajectory, as  $\mathbf{S} = I\boldsymbol{\omega}$  described earlier. The existence of spin of

#### 5.1. ANGULAR MOMENTUM AND SPIN

elementary particles can be established in relativistic quantum mechanics. In non-relativistic quantum mechanics which we study here, spin is introduced as a postulate. Using the algebraic forms (5.4) for extrinsic angular momentum as a guide, we postulate that every elementary particle in a quantum system has an intrinsic spin  $\mathbf{S}$  that satisfies.

$$[S_1, S_2] = i\hbar S_3, \ [S_2, S_1] = i\hbar S_3, \ [S_3, S_1] = \hbar iS_2 \tag{5.10}$$

and

$$S^2 q_{sm} = \hbar s (1+s) q_{sm}, \ S_3 q_{sm} = \hbar m q_{sm}$$
 (5.11)

The eigenvectors  $q_{sm}$  are not functions of  $(r, \theta, \phi)$ ; rather, they are quantities categorized by the half-integers values of indices s and m:

$$s = 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots; \quad m = -s, -s + 1, \cdots, s$$

Every elementary particle has a specific value s of spin, s depending only on the type of particle: for pi mesons, s = 0; for electrons,  $s = \frac{1}{2}$ ; for photons, s = 1, for deltas,  $s = \frac{3}{2}$ ; for gravitons, s = 2; etc. But the spin value s if fixed for any particular particle type.

It follows that the cases of interest here are those for which s = 1/2, which is the spin of particles that make up ordinary matter: electrons, protons, neutrons. In this case, there are just two eigenstates:  $q_{1/2,1/2}$  called *spin up* and symbolically represented by  $\uparrow$ , and  $q_{1/2,-1/2}$ , called *spin down* and represented by  $\downarrow$ . Using these as basis vectors, the state of an s = 1/2 spin particle is the linear combination

$$\mathbf{X} = a\mathbf{X}_{+} + b\mathbf{X}_{-}, \mathbf{X}_{+} = \begin{bmatrix} 1\\0 \end{bmatrix}, \ \mathbf{X}_{-} = \begin{bmatrix} 0\\1 \end{bmatrix}$$
(5.12)

The corresponding eigenvalues are  $+\hbar/2$  for spin up and  $-\hbar/2$  for spin down. The spin operators in (5.10) and (5.11) can now assume the form,

$$\mathbf{S}_1 = \frac{\hbar}{2}\boldsymbol{\sigma}_1, \ \mathbf{S}_2 = \frac{\hbar}{2}\boldsymbol{\sigma}_2, \ \mathbf{S}_3 = \frac{\hbar}{2}\boldsymbol{\sigma}_3$$
 (5.13)

where  $\boldsymbol{\sigma}_i$  are the Pauli spin matrices:

$$\boldsymbol{\sigma}_{1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ \boldsymbol{\sigma}_{2} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \ \boldsymbol{\sigma}_{3} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$
(5.14)

If we set  $\mathbf{S}_{\pm} = \mathbf{S}_1 \pm i\mathbf{S}_2$ , then  $\mathbf{S}_1 = (\mathbf{S}_+ + \mathbf{S}_-)/2$ ,  $\mathbf{S}_2 = (\mathbf{S}_+ - \mathbf{S}_-)/2i$ , and the eigenvectors  $X_{\pm}$  of (5.12), called <u>eigenspins</u> or <u>spinors</u>, satisfy

$$\mathbf{S}_1 \mathbf{X}_+ = \frac{\hbar}{2} \mathbf{X}_- \; ; \; \mathbf{S}_1 \mathbf{X}_- = \frac{\hbar}{2} \mathbf{X}_+,$$
$$\mathbf{S}_2 \mathbf{X}_+ = -\frac{\hbar}{2} \mathbf{X}_- \; ; \; \mathbf{S}_2 \mathbf{X}_- = \frac{\hbar}{2} \mathbf{X}_+$$

# CHAPTER 5. FROM APPROXIMATE THEORIES TO MOLECULAR DYNAMICS

Thus,

$$\mathbf{S}^{2} = \mathbf{S}_{1}^{2} + \mathbf{S}_{2}^{2} + \mathbf{S}_{3}^{2} = rac{3}{4}\hbar^{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{S}_{3} = rac{\hbar}{2} \boldsymbol{\sigma}_{3}$$

and the eigenspinors  $\mathbf{X}_{\pm}$  are clearly eigenvectors of  $\mathbf{S}_3$  with eigenvalues  $\pm \hbar/2$ .

The question that arises at this point is how should electron spin be incorporated into solutions of the Schrödinger equation? This is typically accomplished by writing each one-electron wave function as the product of spatial orbital of the type discussed earlier, and a spin function that depends on the spin of the electron. The electron spin functions have a value of 0 or 1 depending on the quantum number  $m_s$  and are labeled  $\alpha$  and  $\beta$ , with  $\alpha(\frac{1}{2}) = 1$ ,  $\alpha(-\frac{1}{2}) = 0$ ,  $\beta(\frac{1}{2}) = 0$ ,  $\beta(-\frac{1}{2}) = 1$ . In general, the angular spin momentum in an orbital is balanced in the sense that the spin up momentum is balanced with a spin down. The result is that, in general, each spatial orbital can accommodate only (up to) two electrons with paired spins. This principle provides a fundamental rule for describing the electronic structure of polyelectron atoms. We return to this structure later.

As a final remark in this section, we recall from chapter 1 that a spinning charged elementary particle creates a magnetic dipole with a magnetic dipole moment  $\boldsymbol{\mu}$  proportional to the angular spin momentum **S** (recall (1.19)). When placed in a magnetic field **B**, the particle experiences a torque  $\boldsymbol{\mu} \times \mathbf{B}$  with energy  $-\boldsymbol{\mu} \cdot \mathbf{B}$ . The constant  $m_0$  (= -e/m) of proportionality,  $\boldsymbol{\mu} = m_0 \mathbf{S}$ , is called the gyromagnetic ratio, and the Hamiltonian for a spinning particle at rest is  $H = -m_0 \mathbf{B} \cdot \mathbf{S}$ .

### 5.2 Identical Particles and the Pauli Principle

When we extend the ideas covered thus far to multielectron systems, we immediately encounter a question that has profound implications on the form and properties of the wave function. Consider a system of only two perfectly identical particles, particle 1 with coordinates  $\mathbf{r}_1$  and particle 2 with coordinates  $\mathbf{r}_2$ . Ignoring for the moment spin, the wave function of this system is of the form

$$\Psi = \Psi(\mathbf{r}_1, \mathbf{r}_2, t)$$

with Hamiltonian,

$$H = -\frac{\hbar^2}{2m} \Delta_1 - \frac{\hbar^2}{2} \Delta_2 + V(\mathbf{r}_1, \mathbf{r}_2, t)$$

If  $\Psi = \psi(\mathbf{r}_1, \mathbf{r}_2)e^{-iEt/\hbar}$ , the spatial wave function  $\psi$  satisfies the time-independent Schrödinger equation,

$$H\psi = E\psi$$

If particle 1 is in a state  $\psi_1(\mathbf{r}_1)$  and particle 2 is in a state  $\psi_2(\mathbf{r}_2)$ , then

$$\psi(\mathbf{r}_1,\mathbf{r}_2) = \psi(\mathbf{r}_1)\psi_2(\mathbf{r}_2)$$

But if we interchange the two particles (place 2 at location  $\mathbf{r}_1$  and 1 at location  $\mathbf{r}_2$ ), we must arrive at the same dynamical system. We must, in other words, describe the wave function in a way that is noncommittal to which particle is in which state. Either of the following two combinations fulfill the requirement:

$$\psi_{\pm}(\mathbf{r}_1, \mathbf{r}_2) = c(\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) \pm \psi_2(\mathbf{r}_1)\psi_2(\mathbf{r}_2))$$
(5.15)

Thus, quantum mechanics admits two kinds of identical particles, one for which the sign is positive (called *bosons*) and one in which the sign is negative (called *fermions*).

A more concise way to differentiate between the two possibilities is to introduce the *permutation operator* 

$$P_{12}\psi(\mathbf{r}_1,\mathbf{r}_2) = \psi(\mathbf{r}_2,\mathbf{r}_1) \tag{5.16}$$

This operator has the following fundamental properties:

- 1.  $P_{12} = P_{12}^2$  ( $P_{12}$  is a projection)
- 2. The eigenvalues of  $P_{12}$  are  $\pm 1$
- 3.  $P_{12}$  commutes with the Hamiltonian H

$$[H, P_{12}] = 0$$

Thus, P and H are compatible observables and we can find a set of functions that are simultaneous eigenstates of both operators. The wave functions can therefore be either symmetric or antisymmetric with respect to permutation of any two identical particles. The observables  $|\psi|^2$  and  $|P_{12}\psi|^2 = |\psi|^2$  are the same physically possible quantities in either case.

These considerations and experimental evidence lead us to the *Pauli principle*:

Every elementary particle has an intrinsic angular momentum called its spin. Those with integer spins are called bosons, and the corresponding wave function is symmetric with respect to permutations of identical particles ( $P_{12}\psi = \psi$ ). Those with half-integer spins are called fermions, and the corresponding wave function is antisymmetric with respect to permutations of identical particles ( $P_{12}\psi = -\psi$ ).

# CHAPTER 5. FROM APPROXIMATE THEORIES TO MOLECULAR DYNAMICS

Returning to (5.15), if one has two identical fermions,  $\psi_{-}(\mathbf{r}_{1}, \mathbf{r}_{2}) = C(\psi_{1}(\mathbf{r}_{1})\psi_{2}(\mathbf{r}_{2}) - \psi_{1}(\mathbf{r}_{1})\psi_{2}(\mathbf{r}_{2})) = 0$ , a contradiction. Thus, two identical fermions cannot occupy the same state. This observation is referred to as *Pauli's Exclusion Principle*.

We recall that pi mesons and photons are examples of elementary particles with integer spins, while electrons and protons have half-integer spins. Thus, electrons and protons, on which we focus this discussion, are fermions and have antisymmetric wave functions under permutations of identical particles. In the case of n identical particles, we have the permutation operator,

$$P_{123\cdot n}\psi(\mathbf{r}_1,\mathbf{r}_2,\cdot,\mathbf{r}_n)$$
  
=  $\varepsilon_{i_1i_2\cdots i_n}\psi(\mathbf{r}_1,\mathbf{r}_2,\cdot,\mathbf{r}_n)$ 

where  $\varepsilon_{i_1i_2\cdots i_n}$  is the permutation symbol, equal to +1 or -1 if, respectively,  $i_1, i_2, \cdots, i_n$  is an even (symmetric) or odd (antisymmetric) permutation of  $123\cdots n$  and zero if otherwise.

## 5.3 The Helium Atom

The neat level of complexity beyond the hydrogen atom is the helium atom, which consists of two electrons, labeled 1 and 2, with positions  $\mathbf{r}_1$  and  $\mathbf{r}_2$  in space, relative to the origin at the nucleus. The Hamiltonian is

$$H = -\frac{\hbar^2}{2m} \triangle_1 - \frac{\hbar^2}{2m} \triangle_2 - \frac{Ze^2}{4\pi\varepsilon_0 r_1} - \frac{Ze^2}{4\pi\varepsilon_0 r_2} + \frac{e^2}{4\pi\varepsilon_0 r_{12}}$$
(5.17)

where Ze is the charge of the nucleus and  $r_{12} = |\mathbf{r}_2 - \mathbf{r}_1|$ . Schrödinger's equation,

$$H\psi = E\psi$$

is not separable due to the presence of the electron-electron repulsive term involving  $1/r_{12}$ .

Unfortunately, for this system and all others excluding the hydrogen atom, no exact solutions exist. We must therefore resort to approximations. Among the most powerful and successful approximation method is the Ritz method, known in quantum mechanics literature as simply "the variational method". We give a brief account of this method in the next section. The idea is simple: minimize the energy E over a class of *admissible functions* in which the wave function belongs. It is this latter consideration, the identification of admissible functions, that now presents a formidable problem, for admissible functions must be possible spin oribtals or combinations of spin orbitals that are  $L^2(\mathbb{R}^3)$  and, importantly, satisfy the antisymmetric requirement of fermions. The process of seeking such

#### 5.3. THE HELIUM ATOM

admissible functions gives us license to explore various approximations of the Schrödinger equation for the helium atom.

One common approximation of (5.17) that leads to an illustration of how to cope with the antisymmetric issue results when we ignore the  $1/r_{12}$  term and consider the separable Hamiltonian,

$$(H_1 + H_2)\psi(\mathbf{r}_1, \mathbf{r}_2) = E\psi(\mathbf{r}_1, \mathbf{r}_2)$$

where

$$H_i = -\frac{\hbar^2}{2m} \triangle_i - \frac{Ze^2}{4\pi\varepsilon_0 r_i} \quad i = 1, 2$$

For this separable approximation, the wave function can be written as the product of the individual one-electron wave functions,  $\psi_1(\mathbf{r}_1)$  and  $\psi_2(\mathbf{r}_2)$ . Assuming these are normalized, the total energy E is easily found to be the sum of the individual orbital energies,  $E_i = \langle \psi_1, H_i \psi_i \rangle$ , i = 1, 2. Let us denote the 1s orbitals of the hydrogen atom  $\varphi_{1s}(\mathbf{r})$  (recall Table 4.1). Then for a system of two electrons, one might suspect that the wave function

$$\varphi_{1s}(1)\varphi_{1s}(2) \equiv \varphi_{1s}(\mathbf{r}_1)\varphi_{1s}(\mathbf{r}_2) \tag{5.18}$$

be admissible as it has the lowest possible energy state for the system. This satisfies the "indistinguishability" criterion, because it is unchanged when  $\mathbf{r}_1$  is interchanged with  $\mathbf{r}_2$ , and it leads to an energy twice that of a single electron. If we consider the wave function  $\varphi_{2s}$  the 2s-orbitals, the possible wave functions are

$$\varphi_{1s}(1)\varphi_{2s}(2)$$
 and  $\varphi_{1s}(2)\varphi_{2s}(1)$ 

But these functions do not satisfy the indistinguishability criterion: we do not get the same function by interchanging 1 and 2. However, the following linear combinations do satisfy this criteria:

$$(\varphi_{1s}(1)\varphi_{2s}(2) \pm \varphi_{1s}(2)\varphi_{2s}(1))/\sqrt{2} = \varphi_{\pm}^{1;2}$$
(5.19)

The three spatial functions in (5.18) and (5.19) do satisfy the criterion that they are unchanged when electrons 1 and 2 are interchanged, but they are not all antisymmetric; indeed, only  $\varphi_{-}^{1,2}$  is antisymmetric:  $\varphi_{-}^{1,2}(1,2) = -\varphi_{-}^{12}(2,1)$ . But we have yet to take into account spin.

Recall that  $\alpha$  and  $\beta$  are the electron spin functions with  $\alpha(\frac{1}{2}) = 1$ ,  $\alpha(-\frac{1}{2}) = 0$ ,  $\beta(\frac{1}{2}) = 0$ ,  $\beta(-\frac{1}{2}) = 1$ , depending on the quantum number of the electron. For the two electron system, there are four spin states;  $\alpha(1), \beta(1), \alpha(2), \beta(2)$ , with  $\alpha$  denoting the spin state corresponding to spin-quantum number m = 1/2 and  $\beta$  corresponding to m = -1/2. The combined states for two electrons may be

$$\alpha(1)\alpha(2), \beta(1)\beta(2), (\alpha(1)\beta(2) \pm \alpha(2)\beta(1))/\sqrt{2}$$
 (5.20)

# CHAPTER 5. FROM APPROXIMATE THEORIES TO MOLECULAR DYNAMICS

(the  $1/\sqrt{2}$  for normalization). All of these are symmetric with respect to an exchange of electrons except  $\alpha(1)\beta(2) - \alpha(2)\beta(1)/\sqrt{2}$ , which is antisymmetric.

But the Pauli principle must be applied to the combined orbitals and spins: the *spin orbitals*. Consider as possibilities of the product of the spins (5.20) and the orbitals in (5.19). Only the states

$$\varphi_{+}^{1,2}\sigma_{-}, \varphi_{-}^{1,2}\sigma_{+}, \varphi_{-}^{1,2}\alpha(1)\alpha(2), \text{ and } \varphi_{+}^{1,2}\beta(1)\beta(2)$$

where  $\sigma_{\pm} = (\alpha(1)\beta(2) \pm \alpha(2)\beta(1))/\sqrt{2}$ , are antisymmetric, and therefore admissible as spin orbitals for electrons.

In general, wave functions that satisfy the Pauli principle can be written in the form of a determinant called the *Slater determinant*, which exploits the property of determinants that the sign of a determinant is changed when any pair of rows are interchanged. Thus, for two electrons in the ground (1s) state of helium, the wave function

$$\psi(1,2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_{1s}(\mathbf{r}_1)\alpha(1) & \varphi_{12}(\mathbf{r}_2)\alpha(2) \\ \varphi_{1s}(\mathbf{r}_1)\beta(1) & \varphi_{1s}(\mathbf{r}_2)\beta(2) \end{vmatrix}$$
$$= \frac{1}{\sqrt{2}}\varphi_{1s}(\mathbf{r}_1)\varphi_{1s}(\mathbf{r}_2) \left[\alpha(1)\beta(2) - \alpha(2)\beta(2)\right]$$
$$= -\psi(2,1)$$

The above determinant is the Slater determinant for this choice of wave function. The function

$$X_{1s}^{\alpha}(\mathbf{r}_1) = \psi_{1s}(\mathbf{r}_1)\alpha_1(1), \ X_{1s}^{\beta}(\mathbf{r}_1) = \varphi(\mathbf{r}_1)\beta(1)$$

are spin orbitals, and  $\psi(1,2)$  can be expressed as

$$\psi(1,2) = \frac{1}{\sqrt{2}} \begin{vmatrix} X_{1s}^{\alpha}(\mathbf{r}_1) & X_{1s}^{\beta}(\mathbf{r}_1) \\ X_{1s}^{\alpha}(\mathbf{r}_2) & X_{1s}^{\beta}(\mathbf{r}_2) \end{vmatrix}$$

The Slater determinant can be calculated for approximations of wave functions for atoms with any number N of electrons. Denoting the spin orbitals

$$\varphi_{1s}(\mathbf{r}_1)\alpha(1) = 1s(1) \text{ and } \varphi_{1s}(\mathbf{r}_1)\beta(1) = 1s(1)$$

the Slater determinant for Beryllium (with 4 electrons, two in the 1s orbital and two in the 2s orbital),

$$\psi(\mathbf{r}_{1}, \mathbf{r}_{2}, \mathbf{r}_{3}, \mathbf{r}_{4}) \approx \frac{1}{\sqrt{4!}} \begin{vmatrix} 1s(1) & \overline{1s(1)} & 2s(1) & \overline{2s(1)} \\ 1s(2) & \overline{1s(2)} & 2s(2) & \overline{2s(2)} \\ 1s(3) & \overline{1s(3)} & 2s(3) & \overline{2s(3)} \\ 1s(4) & \overline{1s(4)} & 2s(4) & \overline{2s(4)} \end{vmatrix}$$

#### 5.4. VARIATIONAL PRINCIPLE

If any two rows are the same, the determinant has a value 0, so such a state cannot exist by the Pauli principle. Any two electrons tend to avoid each other if they are described by antisymmetric wave functions. In effect, parallel electron spins are avoided. We infer that a maximum of two electrons can occur in an atomic orbital, each with different spins. The electronic structure of an atom consists of a sequence of shells of electron density, each shell consists of the orbitals of a given quantum number n. It is customary to refer to the shell n = 1 as the K-shell, n = 2 the L-shell, n = 3 the M-shell, etc. These properties are fundamental in organizing the various atomic structure of the elements in the periodic table.

### 5.4 Variational Principle

The various wave functions and spin orbitals described earlier must be understood to be only approximations of wave functions for multielectron atoms. They can be used, however, to build good approximations of energy states of atoms using the Rayleigh-Ritz method. To understand the basic idea, we give the following variational theorem.

Let H be the Hamiltonian of a quantum system and let  $E_g$  denote its groundstate energy. Then

$$E_g \leq \langle \psi, H\psi \rangle \quad \forall \psi \in L^2(\mathbb{R}^N), \langle \psi, \psi \rangle = 1$$
(5.21)

In other words,

$$E_g \le \langle H \rangle \tag{5.22}$$

*Proof.* Let  $\psi = \sum_k c_k \psi_k$ ,  $\langle \psi_k, \psi_j \rangle = \delta_{ij}$ . Recall that  $H\psi_k = E_k \psi_k$ . Then

# CHAPTER 5. FROM APPROXIMATE THEORIES TO MOLECULAR DYNAMICS

This variational principle provides a basis for Rayleigh-Ritz approximations of the ground-state energies for more complex atoms. A classical example is the helium atom: a nucleus containing two protons (and some neutrons) and two orbiting electrons. The Hamiltonian is

$$H = -\frac{\hbar^2}{2m}(\Delta_1 + \Delta_2) - \frac{e^2}{4\pi\varepsilon_0} \left(\frac{2}{r_1} + \frac{2}{r_2} - \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\right)$$
(5.23)

In the Ritz method, we introduce a trial function,

$$\psi_a(\mathbf{r}_1, \mathbf{r}_2) = \frac{a^{-3}}{\pi} e^{-(r_1 + r_2)} / a \tag{5.24}$$

where a is a variational parameter. The function  $\psi_a$  is suggested by perturbation theory. The corresponding energy is

$$E(a) = \langle \psi_a, H\psi_a \rangle \tag{5.25}$$

We chose a so that E(a) is a minimum and obtain

$$E_a \approx -2\left(Z - \frac{5}{16}\right)^2 E_H \tag{5.26}$$

where  $Z = 1/2\pi\varepsilon_0$  and  $E_H$  the ground energy of the hydrogen atom [13, page 771]. Numerically,

$$E_a = -76.6eV$$

which compares remarkably well to the experimentally determined value of  $-78.6eV = E_{exp}$ . Note that  $E_a > E_{exp}$ .

## 5.5 Summary of Atomic Structure: The Periodic Table

Let us sum up the properties of atomic structure described up to now and see how they are categorized in making up the Periodic Table of classical chemistry.

All elements of nature are categorized by their atomic structure; the number of electrons and their atomic mass, and these are arranged in a convenient (indeed, an ingenious) way in tabular form, called the Periodic Table, shown in (see Fig. 5.5).

The rows in the table label the *period* of the element which is defined as the number of the energy levels (shells). Thus, elements in row (period) one have



Figure 5.1: The Periodic Table

one energy shell, those in row two have two shells, etc. Thus, there are seven (K, L, M, N, O, P, Q) rows in the table. The columns of the table are called *groups*. In general, elements in a group have the same number of electrons in their outer orbit, but there are exceptions.

Helium, for example, has two electrons in its outer shell, but is listed in group 8. Sodium, for example, is in period 3 (three shells), group 1 (one electron in its outer shell), and has atomic number 11 (11 electrons, 2 in shell k, 8 in shell  $\ell$ , and only 1 in shell m). Chlorine is in period 3, group 17: 3 shells, 17 electrons, all in its *m*-shell, which can hold up to 18. Thus, when sodium and chlorine are brought together, sodium contributes its extra outer electron to the chlorine atom, leaving it with 10 inside its two filled shells and one in the outer shell of the chlorine atom which then acquires the 18 electrons in shell m. The result is a sodium ion, Na+, with one less electron than its atomic number. The number of positive or negative charges an ion possesses is called its *valence* (or *electrovalence*). In this example, Na+ has a positive valance, or a valence of +1.

Atomic bond is a term describing the tendency of atoms and ions to combine in such a way as to produce a stable system with lowest energy (?). Beginning with elements 1-18 in the Periodic Table, the 2-8-8 rule applies: up to 2, then 8, then 8 electrons in shells k,  $\ell$ , and m. Some elements, such as sodium and magnesium, while normal (neutral) atoms, are very unstable and seek a lower energy level by giving up electrons, thereby filling their inner shells, or by acquiring 8 electrons to fill their outer shell. The former is the most likely possibility. Conversely, some atoms are missing electrons in their outer shell (e.g. oxygen and fluorine). Symbolically the situation is illustrated in Fig. 5.2.

To reach a stable state, such atoms can acquire or lose electrons in two ways

CHAPTER 5. FROM APPROXIMATE THEORIES TO MOLECULAR DYNAMICS



Figure 5.2: The simplistic diagrams ignore orbital structures



Figure 5.3: Figure 5.3

1. "*covalent bonding* is an intermolecular form of chemical bonding characterized by the sharing of one or more pairs of electrons between two species, producing a mutual attraction that holds the molecule together" [2]. A more detailed description follows.

2. *Electrovalent* (or *ionic*) *bonding* is an intermolecular chemical bonding in which an ion from one species is transferred to another to create a stable compound molecule.


Figure 5.4: Figure 5.4



Figure 5.5: Figure 5.5

Examples

1. Sodium Oxide

$$\begin{array}{rl} \mathrm{Na}\ 2,8,1 & \mathrm{Na}^{+}\ 2,8 \\ & \Rightarrow & \\ \mathrm{Cl}\ 2,8,7 & \mathrm{Cl}^{-}\ 2,8,8 \end{array} \Rightarrow \mathrm{NaCl} \end{array}$$

2. Magnesium Oxide

$$\begin{array}{rcl} \mathrm{Mg}\ 2,8,2 & \mathrm{Mg}^{2+}\ 2,8 \\ & \Rightarrow & \\ \mathrm{O}\ 2,6, & \mathrm{O}^{2-}\ 2,8 \end{array} \Rightarrow \& \mathrm{MgO} \end{array}$$

2. Calcium Chloride

Cl 2, 8, 7	$Cl^{-} 2, 8, 8$
Ca 2, 8, 8, 2	$\Rightarrow$ Ca 2, 8, 8 $\Rightarrow$ MgO
Cl 2, 8, 7	$Cl^{-} 2, 8, 8$

UHAPTER 0	HAPTER	6
-----------	--------	---

# The Transition from Quantum Mechanics to Approximate Theories and to Molecular Dynamics

### 6.1 Introduction

The enormous complexity of Schrödinger's equations for systems of atoms or molecules has led to the development of numerous approximate theories which have been used very effectively for important applications in chemistry, biology, and materials science. Among methods far removed from quantum theory but applicable to large systems of atoms are those that fall under the classification of models of *molecular dynamics* (MD). While applicable, in principle, to systems involving large numbers of atoms and molecules, MD is limited by many simplifying assumptions that can affect the accuracy of its predictions. Between MD and quantum theory are several more accurate approximate theories that result from various assumptions and conditions placed on the quantum dynamical system. In this chapter, we develop a brief introduction to MD and to various approximate theories.

### 6.2 Molecular Dynamics

Molecular dynamics refers to a class of mathematical models of systems of atoms or molecules in which

1. each atom (or molecule) is represented by a material point in  $\mathbb{R}^3$  assigned

a point mass;

- 2. the motion of the system of mass points is determined by Newtonian mechanics; and
- 3. generally it is assumed that no mass is transferred in or out of the system.

Molecular dynamics (MD) is generally used to calculate ensemble averages of thermochemical and thermomechanical properties of physical systems representing gases, liquids, or solids.

Thus, in MD we consider a collection of N discrete points in  $\mathbb{R}^3$  representing the atom or molecule sites in some bounded domain  $\Omega \subset \mathbb{R}^3$ , each assigned a point mass  $m_i$  and each located relative to a fixed origin **O** by a position vector  $\mathbf{r}_i$ ,  $i = 1, 2, \dots, N$ . The motion of each such particle is assumed to be governed by Newton's second law:

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i \quad , \quad i = 1, 2, \cdots, N$$
(6.1)

where  $\mathbf{F}_i$  is the net force acting on particle *i* representing the interactions of *i* with other particles in the system. These interatomic forces are always assumed to be conservative; i.e. there exists a potential energy  $V = V(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)$  such that

$$\mathbf{F}_{i} = -\frac{\partial V}{\partial \mathbf{r}_{i}} \quad , i = 1, 2, \cdots, N$$
(6.2)

so that finally

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} + \frac{\partial V}{\partial \mathbf{r}_i} = \mathbf{0} \quad , i = 1, 2, \cdots, N$$
(6.3)

To (6.3) we must add initial conditions on  $\mathbf{r}_i(0)$  and  $d\mathbf{r}_i(0)/dt$  and boundary conditions on the boundary  $\partial\Omega$  of  $\Omega$ , these generally chosen to represent a periodic pattern of the motion of the system throughout  $\mathbb{R}^3$ . The solutions  $\{\mathbf{r}_i(t)\}_{i=1}^N$ (6.3) for  $t \in [0, T]$  determine the dynamics of the system. Once these are known, other physical properties of the system such as ensemble averages represented by functionals on the  $\mathbf{r}_i$  and  $d\mathbf{r}_i/dt$  can be calculated. The characterization of the behavior of the dynamical system must also be invariant under changes in the inertial coordinate system and the time frame of reference.

As in quantum systems, it is also possible to describe the equations of MD in the phase space of position-momentum pairs  $(\mathbf{q}, \mathbf{p})$  and in terms of the Hamiltonian of the system,

$$H(\mathbf{r}_{1}, \mathbf{r}_{2}, \cdots, \mathbf{r}_{N} ; \mathbf{p}_{1}, \mathbf{p}_{2}, \cdots, \mathbf{p}_{N})$$
  
=  $\sum_{i=1}^{N} \frac{1}{2m_{i}} \mathbf{p}_{i} \cdot \mathbf{p}_{i} + V(\mathbf{r}_{1}, \mathbf{r}_{2}, \cdots, \mathbf{r}_{N})$  (6.4)

#### 6.2. MOLECULAR DYNAMICS

 $(\mathbf{q}_i \equiv \mathbf{r}_i)$ . Then, instead of (6.3), we have the 2N first-order systems,

$$\dot{\mathbf{r}}_{i} = \frac{\partial H}{\partial \mathbf{p}_{i}} , \quad \dot{\mathbf{p}}_{i} = -\frac{\partial H}{\partial \mathbf{r}_{i}}$$

$$i = 1, 2, \cdots, N$$
(6.5)

wherein  $\dot{\mathbf{r}}_i = d\mathbf{r}_i/dt$  and  $\dot{\mathbf{p}}_i = d\mathbf{p}_i/dt$ . Thus, (5.5) represents 6N ordinary differential equations for the 2N-three-vectors  $(\mathbf{r}_1(t), \mathbf{r}_2(t), \cdots, \mathbf{r}_N(t); \mathbf{p}_1(t), \mathbf{p}_2(t), \cdots, \mathbf{p}_N(t))$ .

One of the most challenging, difficult, and critical aspects of MD is the identification of the appropriate potential function V for the atomic-molecular system at hand. The following is often given as a general form of such potentials:

$$V(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) = \sum_{i=1}^N V_1(\mathbf{r}_i) + \sum_{\substack{i,j=1\\j>i}}^N V_2(\mathbf{r}_i, \mathbf{r}_j)$$
$$+ \sum_{\substack{i,j,k=1\\j>i,k>i}}^3 V_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \cdots$$
(6.6)

where  $V_s$  is called the *s*-body potential. The 1-body potential

$$V_1(\mathbf{r}_1) + V_1(\mathbf{r}_2) + \cdots + V_1(\mathbf{r}_N)$$

is the potential energy due to the external force field, the 2-body potential,

$$V_2(\mathbf{r}_1,\mathbf{r}_2)+V_2(\mathbf{r}_1,\mathbf{r}_3)+\cdots+V_2(\mathbf{r}_{N-1},\mathbf{r}_N)$$

represents pair-wise interactions of particles, the 3-body potential three-particle interactions, etc. The *Lennard-Jones* potential,

$$V(\mathbf{r}_{i}, \mathbf{r}_{j}) = 4\varepsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^{6} \right]$$
$$r_{ij} = |\mathbf{r}_{i} - \mathbf{r}_{j}|$$
(6.7)

 $\sigma$  being a constant that depends upon the atomic structure of the atom or molecule at positions  $\mathbf{r}_i, \mathbf{r}_j$ , and  $\varepsilon$  being the potential energy at the minimum of V is a well-known example of a two-body potential.

We remark that computer simulations employing MD models, approximations of potentials are used to simplify the very large computational problems that can arise. For just pair-wise interactions of N particles, for example, the force  $\partial V/\partial \mathbf{r}_i$ 



**Figure 6.1**: The hydrogen molecule  $H_2$  consisting of two nuclei at a and b and electrons at 1 and 2.

involves  $(N^2 - N)/2$  terms. A typical simplification is to introduce a cut-off radius R around each particle and to include only those interactions with neighboring particles inside that radius. The truncated pair-wise potential is then

$$V_{\text{cut-off}} = f(r)V(r) = \begin{cases} V(r) \ , \ r \le R \\ 0 \ , \ r > \mathbb{R} \end{cases}$$

where  $r = |\mathbf{r}_i - \mathbf{r}_j|$  and f is a smooth cut-off function varying from 1 at r = 0 to 0 at r = R.

## 6.3 The Connection Between MD and Quantum Mechanics

We shall now show how the results we derived for simple quantum systems can be used to obtain force potentials for a simplified molecular dynamics model. The issue is one of scale. Quantum effects, we presume, are essentially confined to the individual atom or molecule, while MD treats these as material points-particles.

The ideas can be illustrated through the example of the hydrogen molecule (Cf. Liu et al [12])  $H_2$  consisting of two protons and two electrons. The position of the electrons relative to protons at  $\mathbf{r}_a$  and  $\mathbf{r}_b$  are denoted  $\mathbf{r}_{ai}$ ,  $\mathbf{r}_{bi}$ , i = 1, 2, and  $\mathbf{r}_{12}$  is the distance between electrons. The distance separating the two atoms is denoted r. This situation is illustrated in Fig. 6.1.

The analysis proceeds in the following steps:

1. the total energy E of the system is the sum of the energies of the two unbounded hydrogen atoms  $E_a$  and  $E_b$ , plus the energy due to interactions U, which we call the *binding energy*,

$$E = E_a + E_b + U$$

2. MD assumes that no energy is absorbed or emitted by the atoms, so the energies  $E_a, E_b$  are the ground states calculated in Chapter 4 (recall (4.31)):

$$E_1 = E_a = E_b = -\frac{m}{2\hbar^2} \left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 = -13.6eV$$

3. the energy of the coupled system can be written as a function of the separation distance  $r(|\mathbf{r}_b - \mathbf{r}_a|)$ , now regarded as a real parameter:

$$U(r) = V(r) = E(r) - 2E_1$$

Thus,  $V(r) = V_2(r)$  is the potential function for 2-body (pair-wise) interactions.

4. the protons at  $\mathbf{r}_a$  and  $\mathbf{r}_b$  are regarded as stationary at the distance r, an approximation known as the *Bohr-Oppenheimer* approximation, which is justified on the basis that the mass of the proton is 1,800 times that of the electron.

It remains to calculate the energy E(r). For this we must resort to quantum mechanics. Under the conditions and assumptions listed above, the Hamiltonian is

$$\tilde{H} = -\frac{\hbar^2}{2m} \sum_{i=1}^2 \Delta_i - \frac{1}{4\pi\varepsilon_0} \left\{ \sum_{i=1}^2 \left( \frac{e^2}{r_{ai}} + \frac{e^2}{r_{bi}} \right) + \frac{e^2}{r_{12}} + \frac{e^2}{r} \right\}$$
(6.8)

The corresponding Schrödinger equation is

$$\tilde{H}\Psi = E\Psi 
= (V(r) + 2E_1)\Psi$$
(6.9)

Since

$$E = \langle \Psi, H\Psi \rangle \quad (\langle \Psi, \Psi \rangle = 1)$$

we have

$$V(r) = \langle \Psi, H\Psi \rangle - 2E_1 \tag{6.10}$$

where, we assume, the wave function  $\Psi$  depends parimetrically on r.

In general, we can write the coupled MD-quantum mechanics.

$$\tilde{H}\Psi = \left(V + \sum_{\alpha=1}^{M} E_{\alpha}\right)\Psi$$

$$m_{i}\ddot{\mathbf{r}}_{i} = -\frac{\partial V}{\partial\mathbf{r}_{i}} \quad , \ i = 1, 2, \cdots, N$$
(6.11)

for M-electron systems ( $\ddot{\mathbf{r}}_i = d\dot{\mathbf{r}}_i/dt = d^2 \mathbf{r}_i/dt^2$ ).

It is understandably rare that the wave function  $\Psi$  can be computed exactly. Then we may resort to developing approximations of E such as those described in connection with the Ritz method in Chapter 4. Then, for a given trial function  $\varphi \in (L^2(\mathbb{R}^{3M}))$ , we set

$$E \approx \frac{\langle \varphi, H\varphi \rangle}{\langle \varphi, \varphi \rangle} = E_{\varphi} \tag{6.12}$$

The trial functions are of the form,  $\varphi = \sum_k c_k \chi_k$  where  $\chi_k$  are specific basis functions. The coefficients  $c_k$  are chosen so that  $E_{\varphi}$  is a minimum in the space spanned by the  $\chi_k$ 's:  $\partial E_{\varphi}/\partial c_k = 0$ . The resulting trial function  $\varphi$  is used to commute the minimum  $E_{\varphi^*}$ . Then

$$V \approx E_{\varphi^*} - \sum_{\alpha=1}^M E_\alpha \tag{6.13}$$

We describe other approximate methods in the following sections.

### 6.4 Some Approximate Methods

We briefly outline three of the more popular approximate methods. Two of these are based on Ritz-type approximations similar to those described earlier.

### 6.4.1 The LCAO Method

The following variational approach is called the *tight binding method* or the *linear* combination of atomic orbitals (LCAO) method by Liu et al [12] and is attributed to Bloch [1] and Slater and Koster [16]. A molecular orbital is defined as a wave function for a single electron in a non-central field containing two or more nuclei. For the hydrogen ion  $H_2^+$ , the Hamiltonian is

#### 6.4. SOME APPROXIMATE METHODS

$$H = -\frac{\hbar^2}{2m} \sum_{i=1}^2 \Delta_i - \frac{1}{4\pi\varepsilon_o} \left\{ \sum_{\alpha=1}^2 \left( \frac{e^2}{r_{\alpha 1}} + \frac{e^2}{r_{\alpha 2}} \right) - \frac{e^2}{r} \right\}$$
(6.14)

where we use the notation in Fig. 6.1. Let  $\psi_a$  and  $\psi_b$  be the molecular orbitals for a single hydrogen atom (i.e.  $(-\hbar^2/2m\Delta - e^2/4\pi\varepsilon_0 r_{a1})\psi_a = 0$ ,  $(-\hbar^2/2m\Delta - e^2/4\pi\varepsilon_0 r_{b2})\psi_b = 0$ . Then, as an approximation of the wave function for H of (6.14), we take the linear combination,

$$\Psi = \alpha \psi_a + \beta \psi_b \tag{6.15}$$

The corresponding approximate energy is

$$\tilde{E} = \frac{\left\langle \tilde{\Psi}, H\tilde{\Psi} \right\rangle}{\left\langle \tilde{\Psi}, \tilde{\Psi} \right\rangle} = \frac{\alpha^2 H_{aa} + 2\alpha\beta H_{ab} + \beta^2 H_{bb}}{\alpha^2 + \beta^2 + 2\alpha\beta G_{\alpha\beta}}$$
(6.16)

where

$$H_{ab} = \langle \psi_a, H\psi_b \rangle \quad , \ G_{\alpha\beta} = \langle \psi_a, \psi_b \rangle$$

Choosing  $\alpha$  and  $\beta$  to minimize  $\tilde{E}$ , we have

 $\alpha = + \beta$ 

so that two molecular orbits for  $H_2^+$  are obtained,

$$\psi^+ = A(\psi_a + \psi_b)$$
 and  $\psi^- = B(\psi_a - \psi_b)$  (6.17)

where A and B are normalizing factors. The corresponding energy states are

$$\tilde{E}^{+} = \frac{H_{aa} + H_{ab}}{1 + G_{ab}} \quad \text{and} \quad \tilde{E}^{-} = \frac{H_{aa} - H_{ab}}{1 - G_{ab}}$$
(6.18)

An approximation of the potential V can now be obtained from (6.13).

This approximate approach does not necessarily respect the Pauli exclusion principle (discussed in an appendix), and therefore can be ineffective or require additional information. The Hartree models, particularly Hartree-Fock, are designed to address these issues.

### 6.4.2 The Hartree-Fock Model

The physical system considered is a molecular system made up of M nuclei and N electrons. The goal is to calculate the ground state energy of the system. Such a system is illustrated in Fig. 6.2. To reduce this M + N body problem to one more tractable, we introduce again the Bohr-Oppenheimer approximation, whereby the M nuclei are considered as classical point particles with charges  $Z_k$  at positions  $\mathbf{R}_k$ ,  $k = 1, 2, \dots, M$ . Within the Bohr-Oppenheimer approximation, the problem of determining the ground state reduces to the nested minimization problems:

$$\inf_{\substack{\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_M \\ \in \mathbb{R}^{3M}}} \left\{ E(\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_M) + c \sum_{1 \le k < j \le M} \frac{Z_k Z_j}{R_{kj}} \right\}$$
(6.19)

with  $R_{kj} = |\mathbf{R}_k - R_M|$  and c the appropriate constant, and

$$E(\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_M) = \inf_{\psi \in \mathcal{H}} \{ \langle \psi, H\psi \rangle , \|\psi\| = 1 \}$$
(6.20)

where

$$H = -\sum_{i=1}^{N} \frac{\hbar^2}{2m} \Delta_{\mathbf{r}_i} - c \sum_{i=1}^{N} \left( \sum_{k=1}^{M} \frac{eZ_k}{|\mathbf{r}_k - \mathbf{R}_k|} \right) + c \sum_{1 \le i < j \le N} \frac{e^2}{r_{ij}}$$
(6.21)

Here  $\triangle_{\mathbf{r}_i} = \frac{\partial^2}{\partial r_{i1}^2} + \frac{\partial^2}{\partial r_{i2}^2} + \frac{\partial^2}{\partial r_{i3}^2}$ ,  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$  and  $\mathbf{r}_i$  is the position vector of electron i, and

$$\mathcal{H} = \prod_{i=1}^{N} H^1(\mathbb{R}^3, \mathbb{C}) \tag{6.22}$$

with  $H^1(\mathbb{R}^3, \mathbb{C})$  denoting the Sobolev space of complex-valued functions (equivalence classes of functions) with generalized first derivatives in  $L^2(\mathbb{R})$ . Hereafter, the nuclei positions  $\mathbf{R}_k$  are regarded as parameters, and we focus on the internal electron minimization problem (6.20).

The Hartree-Fock approximation amounts to a variational formulation of (6.20) designed to respect the Pauli exclusion principle, which is discussed in an appendix to this chapter. First, we derive a reduced functional. Since electrons correspond to fermions, we seek antisymmetric functions in H characterized by Slater determinants of the form,

$$\psi = \frac{1}{\sqrt{N!}} det(\varphi_i(\mathbf{r}_j))$$



Figure 6.2: A molecular system consisting of M nuclei and N electrons.

where  $\varphi_i$  are orthonormal molecular orbitals,

$$\langle \varphi_i, \varphi_j \rangle = \delta_{ij} \ , \ i, j = 1, 2, \cdots, N$$

With this constraint, the Hartree-Fock energy functional becomes

$$E^{HF}(\{\varphi_i\}) = \sum_{i=1}^{N} \frac{\hbar^2}{2m} \|\nabla_{\mathbf{r}_i}\varphi_i\|^2 + \int_{\mathbb{R}^3} \varrho_{\varphi} V dr$$
$$+ \frac{c}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\varrho_{\varphi}(\mathbf{r})\varrho_{\varphi}(\mathbf{r}^1)}{|\mathbf{r} - \mathbf{r}'|} dr dr'$$
$$+ \frac{c}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \sum_{i=1}^{N} \varphi_i(\mathbf{r})\varphi_i^*(\mathbf{r}) dr dr'$$
(6.23)

where  $dr = d^N r = dr_1 dr_2 \cdots dr_N$ ,

$$arrho_arphi({f r}) = \sum_{i=1}^N |arphi_i({f r})|^2$$

The Hartree-Fock reduced minimization thus reads,

$$\inf_{\varphi_i} \{ E^{HF}(\{\varphi_i\}) , \quad \varphi_i \in H^1(\mathbb{R}, \mathbb{C}), \quad \langle \varphi_i, \varphi_j \rangle \\
= \delta_{ij} , \quad i, j = 1, 2, \cdots, N \}$$
(6.24)

The orthonormality condition  $\langle \varphi_i, \varphi_j \rangle = \delta_{ij}$  is a constraint on the choices of test functions.

The Hartree-Fock method consists of constructing Ritz approximations of the variational problem (6.24). We construct a finite number of molecular orbits  $\{\chi_m\}, 1 \leq m \leq n$ , (or approximate molecular orbits) and set

$$\varphi_i \approx \varphi_i^{(m)} = \sum_{k=1}^m C_{ki} \chi_k \tag{6.25}$$

where  $C_{ki}$  are constants to be determined so as to minimize  $E^{HF}$  over span  $\{\varphi_i^{(m)}, i = 1, 2, \dots, N; m = 1, 2, \dots, n\}$ . We demand that

$$\delta_{ij} = \left\langle \varphi_i^{(m)}, \varphi_j^{(m)} \right\rangle = C_{kj}^* X_{k\ell} C_{ki}$$

where  $X_{ij} = \langle \chi_i, \chi_j \rangle$ . Subject to this constraint, substitution of (6.25) into (6.23) leads to a reduced *HF* energy functional of the constants  $C_{ki}$ :

$$E^{EF}(\{\varphi_i^{(m)}\}) = \tilde{E}^{EF}(C_{kj})$$
(6.26)

We then compute the specific coefficient  $\hat{C}_{kj}$  such that

$$\frac{\partial \tilde{E}^{EF}}{\partial C_{kj}}(\hat{C}_{kj}) = 0 \tag{6.27}$$

This is a general approximation process behind the Hartree-Fock model. There are many variants of this approach (see, e.g. Cances [3] or Lions [11] for details and additional references).

### 6.4.3 Density Functional Theory

The density functional theory addresses the calculation of ground states of atomic and molecular systems by characterizing the state of the system using the electron density  $\rho(\mathbf{r})$  instead of the wave function  $\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)$ . For any quantum system involving N electrons, the electron density  $\rho(\mathbf{r})$  is defined as the number of electrons per unit volume, and is given in terms of the wave function by (recalling that now  $\Psi$  is invariant under permutation)

$$\varrho(\mathbf{r}_1) = N \int_{\mathbb{R}} |\Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N)|^2 dx_2 dx_3 \cdots dx_N$$
(6.28)

so that with  $\mathbf{r} = (x, y, z) \in \mathbb{R}^3$ 

$$\int_{\mathbb{R}^3} \rho(\mathbf{r}) dr = N \tag{6.29}$$

#### 6.4. SOME APPROXIMATE METHODS

Once again we consider the time-independent Schrödinger equation for a system of M nuclei of charge  $Z_k$  and N electrons with charges e, and we invoke the Bohr-Oppenheimer approximation, assuming that the electrons adjust to any change in the nuclei. The total energy E in the ground state can then be expressed in terms of the electronic density, which can be expressed as a sum of molecular orbitals  $\varphi_i$  for each electron:

$$\varrho = \sum_{i=1}^{N} |\varphi_i|^2 \tag{6.30}$$

For this system, the Hamiltonian operator is of the form

$$\tilde{H} = \tilde{T} + V_{ne} + V_{ee} \tag{6.31}$$

where  $\tilde{T}$  is the kinetic energy operator,

$$\tilde{T} = \frac{\hbar^2}{2m} \sum_{i=1}^{N} \triangle_{\mathbf{r}_i}$$
(6.32)

 $V_{ne}$  is the external potential for electron *i* due to nuclei  $\alpha$ ,

$$V_{ne} = -\sum_{i=1}^{N} \sum_{\alpha=1}^{M} \frac{cZ_{\alpha}e}{|\mathbf{r}_i - \mathbf{R}_{\alpha}|}$$
(6.33)

and

$$V_{ee} = \sum_{\substack{i,j=1\\i< j}}^{N} \frac{ce^2}{|\mathbf{r}_i - \mathbf{r}_j|}$$
(6.34)

with c the usual constraint.

The electronic energy E satisfies, as usual,

$$H\Psi = E\Psi \tag{6.35}$$

and the total energy is

$$W = E + c \sum_{\substack{\alpha,\beta=1\\\alpha<\beta}}^{M} \frac{Z_{\alpha}Z_{\beta}}{R_{\alpha\beta}}$$
(6.36)

 $R_{\alpha\beta} = |\mathbf{R}_{\alpha} - \mathbf{R}_{\beta}|.$ 

Our goal is to determine these various energy components in terms of the density  $\varrho$ .

The Kohn-Sham form of the energy functional [10] is

$$E(\varrho) = T(\varrho) + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\varrho(\mathbf{x})\varrho(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} dx dy + E_x(\varrho) + E_c(\varrho)$$
(6.37)

where  $T(\rho)$  is the kinetic energy functional,

$$T(\varrho) = \inf\left\{\frac{\hbar^2}{2m}\sum_{i=1}^N \int_{\mathbb{R}^3} |\nabla\varphi_i|^2 dx , \quad \varrho = \sum_{i=1}^N |\varphi_i|^2\right\}$$
(6.38)

and  $E_x$  and  $E_c$  are the so-called exchange and correlation energies. Various approximations of  $E_x$  and  $E_c$  can be introduced. The local density approximation (LDA) employs the exchange energy of a homogeneous electron gas,

$$E_x(\varrho) = -c_x \int_{\mathbb{R}^3} \varrho^{4/3} dr \tag{6.39}$$

where  $c_x$  is a positive constant (=  $3(3/\pi)^{1/3}/4$  in atomic units). For  $E_c$ , similar empirical relations can be employed. We then seek a minimum energy subject to the constraints due to the definition of  $\rho$ :

$$\inf\left\{E(\varrho) + \int_{\mathbb{R}^3} \varrho V_{nec} dr \ , \ \varrho \ge 0 \ , \ \int_{\mathbb{R}^3} \varrho dr = N\right\}$$
(6.40)

where  $V_{nec}$  is the electrostatic potential given by (6.6).

Other energy fuctionals of the density have been proposed. The Thomas-Fermi models employ functionals of the type,

$$E^{TF}(\varrho) = c_o \int_{\mathbb{R}^3} |\nabla \sqrt{\varrho}|^2 dr + c_1 \int_{\mathbb{R}^3} \varrho^{5/3} dr + \frac{1}{2} \int_{\mathbb{R}^3} \frac{\varrho(\mathbf{x})\varrho(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} dx dy - \sum_{k=1}^M \int_{\mathbb{R}^3} \frac{\varrho(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_k|} dr + \frac{1}{2} \sum_{\mathbf{x} \neq \mathbf{y}} \frac{ce^2}{|\mathbf{x} - \mathbf{y}|}$$
(6.41)

and we seek,

$$\inf\left\{E^{TF}(\varrho): \ \varrho \ge 0, \ \sqrt{\varrho} \ \in \ H^1(\mathbb{R}^3), \ \int_{\mathbb{R}^3} \varrho dr = N\right\}$$
(6.42)

## Bibliography

- [1] BLOCH, F. "Uber die quantenmechanik der elektron in kristallgittern," Zeitschrift für Physik, 52, pp.522-600, 1928.
- [2] BORN, M. <u>Atomic Physics</u>, Eighth Edition, Revised by R.J. Blin-Stoyle and J.M. Radcliffe. Dover, New York, 1989.
- [3] CANCES, E. "SCF algorithms for HF electronic calculations," <u>Mathematical</u> <u>Models and Methods for Ab Initio Quantum Chemistry</u>, Edited by M. Defranceschi and C. LeBris, Lecture Notes in Chemistry, Springer-Verlag, Berlin, Heidelberg, N.Y., pp.17-33, 2000.
- [4] DEFRANCESCHI, M. AND LEBRIS, C., (EDS.) <u>Mathematical Models and</u> <u>Methods in Ab Initio Quantum Chemistry</u>. Springer-Verlag, Berlin, Heidelberg, New York, 2000.
- [5] DREIZLER, R.M. AND GROSS, E.K.U. <u>Density Function Theory</u>, Springer-Verlag, Berlin, Heidelberg, New York, 1993.
- [6] GRIFFITHS, D.J. <u>Introduction to Quantum Mechanics</u>, Prentice-Hall, Upper Saddle River, NJ, 1995.
- [7] HALLIDAY, D., RESNICK, R., AND WALKER, J. <u>Fundamentals of Physics</u>, Sixth Edition, John Wiley and Sons, Hoboken, NJ, 2001.
- [8] HINCHLIFFE, A. Computational Quantum Chemistry, John Wiley and Sons, Chichester, 1988.
- [9] JACKSON, J.D. <u>Classical Electrodynamics</u>, Second Edition, John Wiley and Sons, New York et., 1962.

- [10] KOHN, W. AND SHAM, L.J. "Self-Consistent Equations Including Exchange and Correlation Effects", Physics Review, 94, P. 111, 1954.
- [11] LIONS, P.L. "Solutions of Hartree-Fock Equations for Coulomb Systems", Communications in Mathematical Physics, 109, pp.33-97, 1987.
- [12] LIU, W.K., KARPOV, E.G., AND PARK, H.S. "An introduction to computational nanomechanics and materials," <u>Computer Methods in Applied</u> Mechanics and Engineering, Vol. 193, pp.1529-1578, 2004.
- [13] MESSIAH, A. <u>Quantum Mechanics</u>, Volumes 1 and 2. Dover, Mineola, NY, 1999.
- [14] PARR, R.G. AND YANG, W. <u>Density Function Theory of Atoms and</u> Molecules, Oxford University Press, New York, 1989.
- [15] RATNER, M.A. AND SCHATZ, G.C. Introduction to Quantum Mechanics in Chemistry, Prentice-Hall, Upper SAddle River, New York, 2001.
- [16] SLATER, V.C., AND KOSTER, G.F. "Simplified LCAO method for periodic potential problem," Physics Review, 94(6), pp.1498-1524, 1954.
- [17] VON NEUMANN, J. <u>Mathematical Foundations of Quantum Mechanics</u>, University Press, Princeton, NJ, 1955.