Lectures on an Introduction to Quantum mechanics

J. Tinsley Oden

ICES
The University of Texas at Austin
2004

Preface

This monograph contains class notes I prepared during the summer of 2004 based on 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. 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.

Thanks go to Frederika Tausch for typing most of the manuscript and to Derek Ploor for typing some of the early sections.

J. Tinsley Oden
February 3, 2005

Contents

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

$$\boxed{F = k\frac{|q_1|\,|q_2|}{r^2}} \tag{1.1}$$

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 ϵ_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,

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

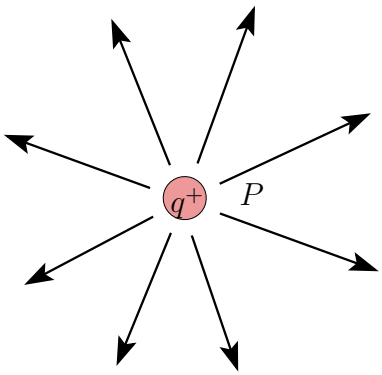
   The fact that electrical charge is "quantized" (meaning discretely defined as an integer multiple of e) is regarded by some as "one of the most profound mysteries of the physical world" (Cf. [9, page 251])

2. charge is *conserved*: the net charge in a system or object is preserved.

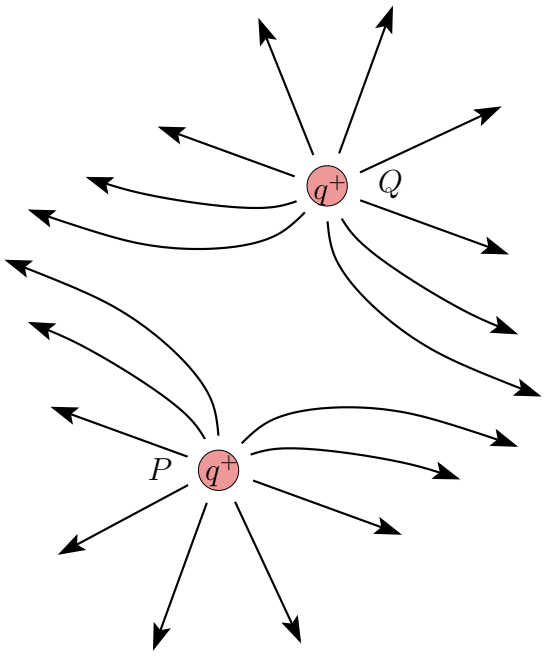
The fundamental notion of an electric field is intimately tied to the force field generated by electrical charges. Let q_1 denote a positive point charge situated at a point P in space. Imagine that a second positive point charge q_2 is placed at point Q near to P. According to Coulomb's law, q_1 exerts 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 $\boldsymbol{E}$ in the space surrounding it, such that the magnitude of $\boldsymbol{E}$ 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

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



(a) Positive charge at point P



(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.

For m such charges,

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

$\boldsymbol{F}_{0i}$ 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.

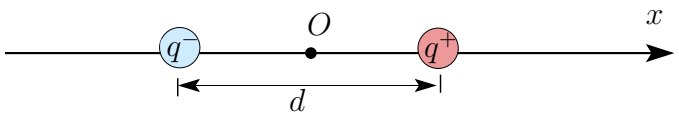


Figure 1.2: Equal and opposite charges on a line

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

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

(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.

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 through the surface.

1.2 Gauss' Law

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

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

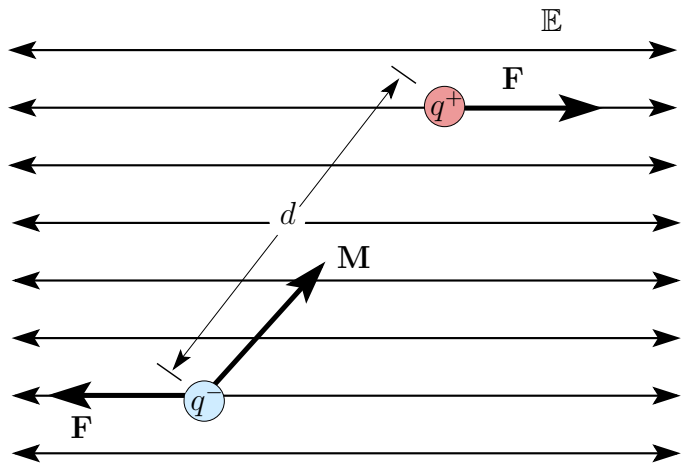


Figure 1.3: Equal and opposite charges in an electric field $\boldsymbol{E}$ a distance d apart

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

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

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

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 ΔU 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

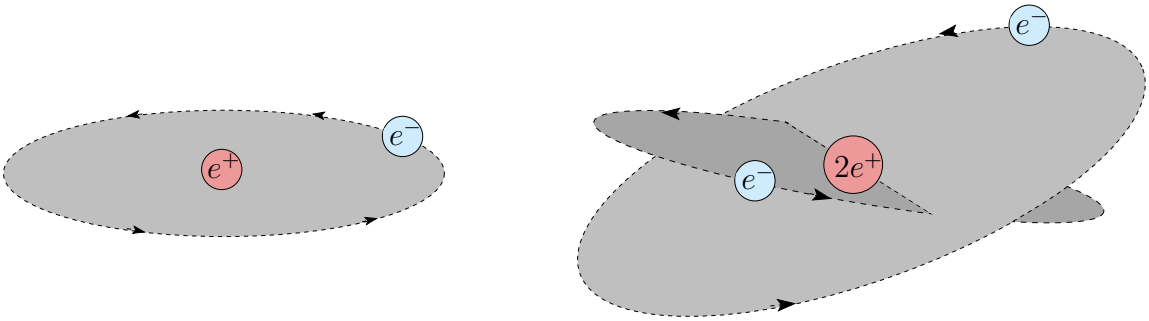


Figure 1.4: Model of an atom as charged electrons in orbits around a nucleus

model was discarded when Bohr introduced the quantum model of atoms in which, for any atom, electrons can only exist in so-called discrete quantum states of well-defined energy or so-called 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.

This is no "monopole" analogy 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^{-4} tesla). Since the motion of electric charge is called *current*, it is easily shown that $\boldsymbol{B}$ can be expressed in terms of the current i for various motions of charges. *Ampere's Law* asserts that on any closed loop $\mathcal{C}$ in a plane,

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

where μ_0 is the *permeability constant*, $\mu_0 = 4\pi \times 10^{-7}$ T-m/A, and i_{enclosed} is the net current flowing perpendicular to the plane in the planar region enclosed by $\mathcal{C}$. 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*,

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

where Φ_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 $\mathcal{C}$.

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

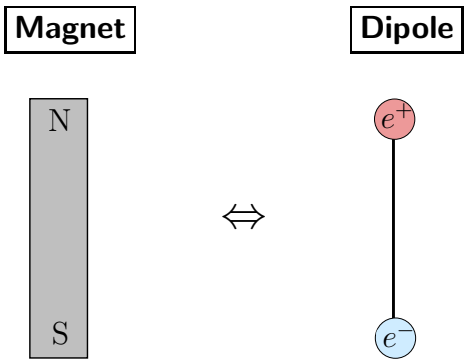


Figure 1.5: Electric dipole analogous to a dipole caused by a magnet

a particle of charge q moving around a closed loop $\mathcal{C}$ encompassing an area A with unit normal $\boldsymbol{n}$. Then $\boldsymbol{B}$ induces an electric field $\boldsymbol{E}$ such that

$$\boxed{\oint_{\mathcal{C}} \boldsymbol{E} \cdot \, \mathrm{d}\boldsymbol{s} = -\frac{\mathrm{d}}{\mathrm{d}t} \int_A \boldsymbol{B} \cdot \boldsymbol{n} \, \mathrm{d}A} \tag{1.7}$$

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 $\boldsymbol{\tau}$, called the *spin angular momentum*, and an intrinsic *spin magnetic dipole moment* $\boldsymbol{\mu}$, related by

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

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

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 evaluation 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 $\boldsymbol{x} = (x_1, x_2, x_3)$ 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,

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

or, for simplicity, of the form

$$\boxed{u(x,t) = u_0 \sin(kx - \omega t)} \tag{1.11}$$

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

Here

$$\left.\begin{array}{rcl} u_0 & = & \text{the amplitude of the wave} \\ k & = & \text{the angular wave number} \\ \omega & = & \text{the angular frequency} \end{array}\right\} \tag{1.12}$$

and

$$\left.\begin{array}{rcccl} \lambda & = & 2\pi/k & = & \text{the wave length} \\ T & = & 2\pi/\omega & = & \text{the period (of oscillation)} \end{array}\right\} \tag{1.13}$$

The frequency ν of the wave is defined as

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

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

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

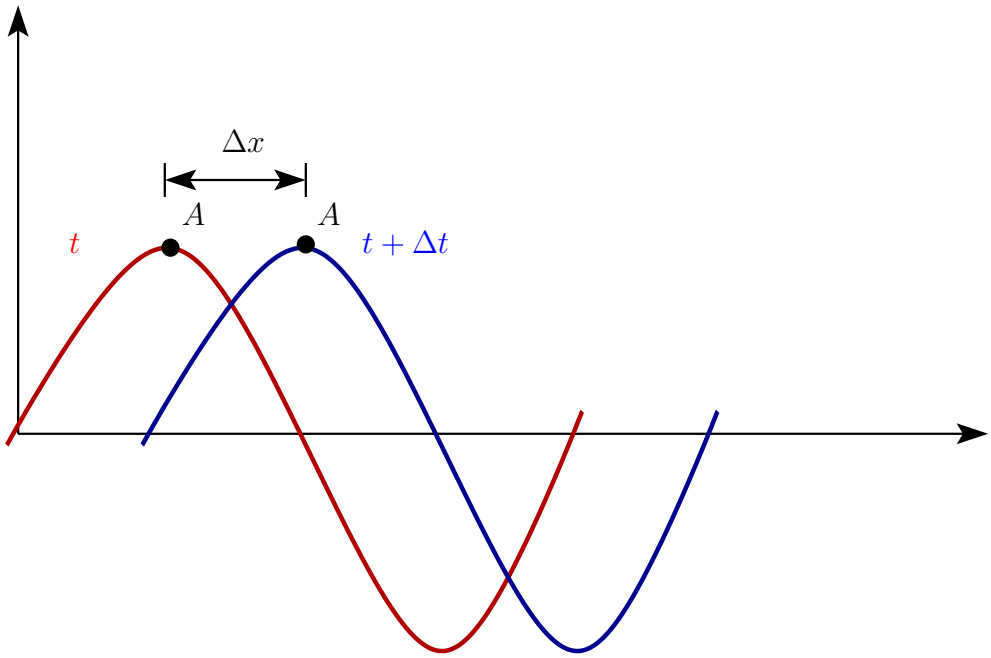


Figure 1.6: Incremental motion of a wave from over a time increment Δt

so the wave speed is

$$\boxed{v = \frac{\omega}{k} = \frac{\lambda}{T} = \lambda\nu} \tag{1.15}$$

The quantity

$$\varphi = kx - \omega t = \text{the phase of the wave} \tag{1.16}$$

Two waves of the form

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

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

$$\begin{aligned} u(x,t) &= u_1(x,t) + u_2(x,t) \\ &= (2u_0 \cos\varphi/2) \sin(kx - \omega t + \varphi/2) \end{aligned} \tag{1.17}$$

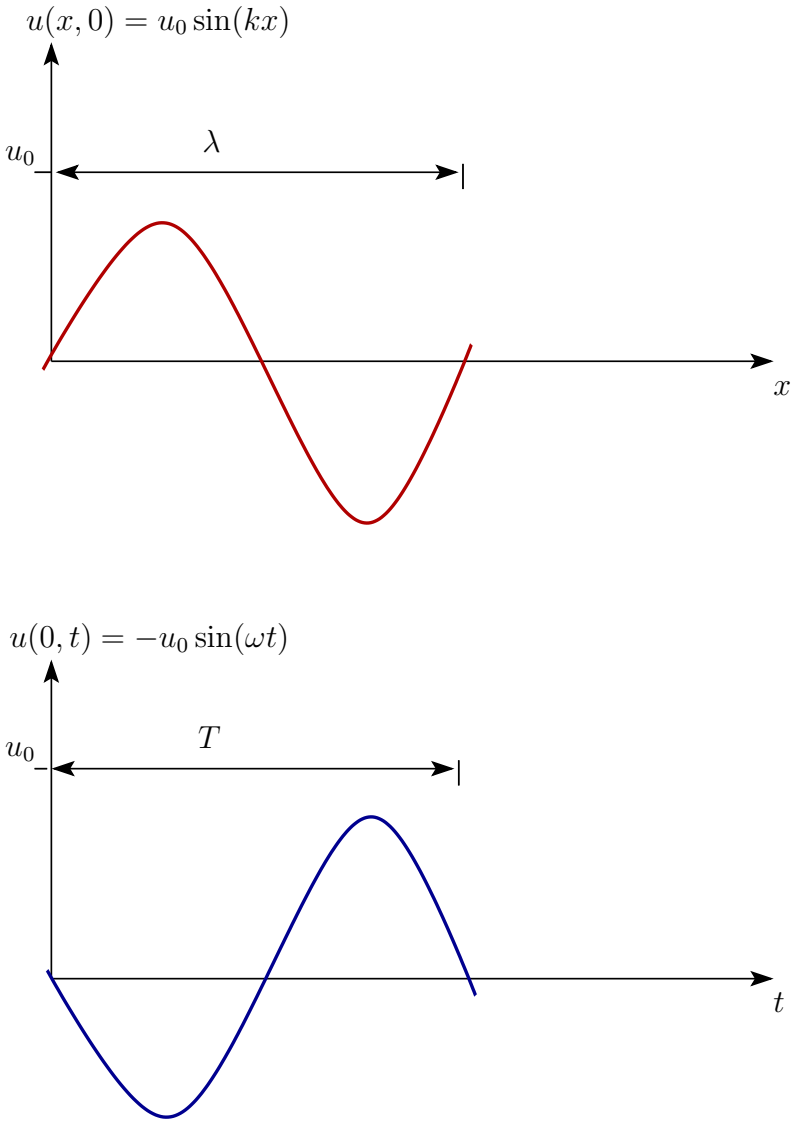


Figure 1.7: Properties of a simple plane wave

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, ω/k is recognized as the wave speed.

1.6 Maxwell's Equations

Consider electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$ in a vacuum. Recall the following physical laws:

Gauss' Law (Conservation of charge in a volume Ω 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 $\mathrm{d}A$ and ρ is the charge density, q_Ω being the total charge contained in Ω.

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

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

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

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

$$\begin{aligned}\int_{\mathcal{C}} \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{s} &= \mu_0 i_{\text{enclosed}} + \mu_0\epsilon_0 \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_A \boldsymbol{E}\cdot\boldsymbol{n}\,\mathrm{d}A\right)\\ &= \mu_0\int_A \boldsymbol{j}\,\mathrm{d}A + \mu_0\epsilon_0\int_A \boldsymbol{n}\cdot\frac{\partial \boldsymbol{E}}{\partial t}\,\mathrm{d}A\end{aligned} \tag{1.21}$$

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

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:

$$\boxed{\begin{aligned}\epsilon_0\boldsymbol{\nabla}\cdot\boldsymbol{E} &= 4\pi\rho\\ \boldsymbol{\nabla}\times\boldsymbol{E} &= -\tfrac{\partial \boldsymbol{B}}{\partial t}\\ \boldsymbol{\nabla}\times\boldsymbol{B} &= \mu_0\boldsymbol{j} + \mu_0\epsilon_0\tfrac{\partial \boldsymbol{E}}{\partial t}\\ \boldsymbol{\nabla}\cdot\boldsymbol{B} &= 0\end{aligned}} \tag{1.23}$$

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

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

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 $\boldsymbol{j}$:

$$\boxed{\begin{aligned} \boldsymbol{\nabla}\cdot\boldsymbol{D} &= \rho_f \\ \boldsymbol{\nabla}\times\boldsymbol{E} &= -\tfrac{\partial \boldsymbol{B}}{\partial t} \\ \boldsymbol{\nabla}\cdot\boldsymbol{B} &= 0 \\ \boldsymbol{\nabla}\times\boldsymbol{H} &= \boldsymbol{j}_f + \tfrac{\partial \boldsymbol{D}}{\partial t} \end{aligned}} \tag{1.25}$$

where ρ_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 $\boldsymbol{j} = \boldsymbol{0}$. Then

$$\frac{\partial}{\partial t}\boldsymbol{\nabla}\times\boldsymbol{B} = \boldsymbol{\nabla}\times\frac{\partial \boldsymbol{B}}{\partial t} = -\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\boldsymbol{E} = \mu_0\epsilon_0\frac{\partial^2\boldsymbol{E}}{\partial t^2}$$

so that we arrive at the wave equation,

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

where the wave speed is

$$c = \frac{1}{\sqrt{\mu_0\epsilon_0}} = 3.0\times10^8\ \mathrm{m/s} \tag{1.27}$$

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

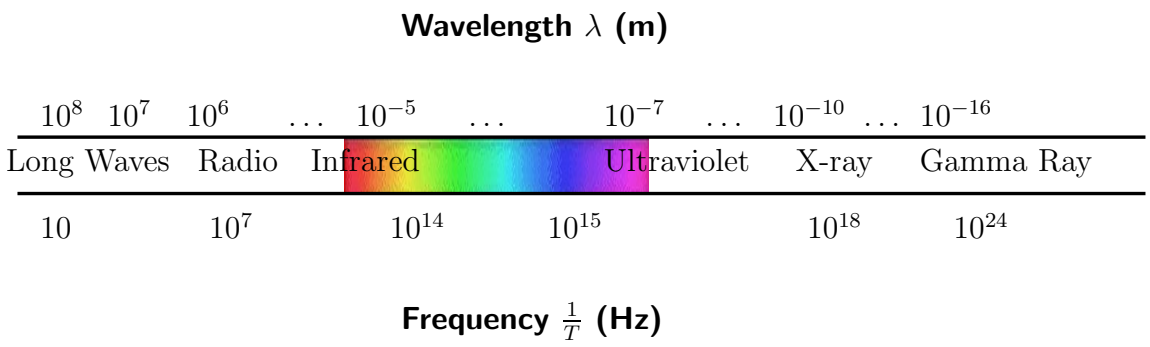


Figure 1.8: Wavelengths and frequencies of electromagnetic waves.

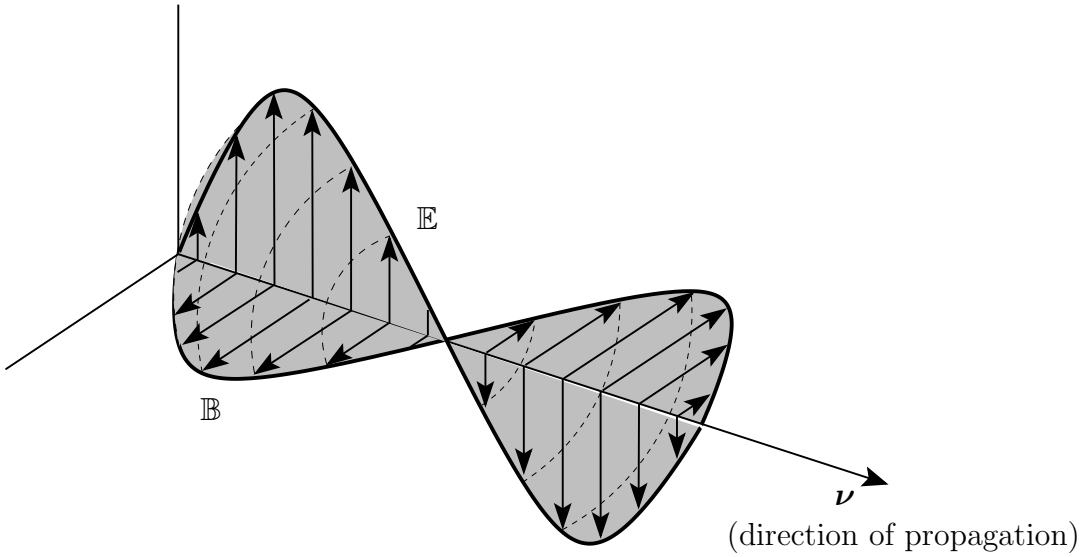


Figure 1.9: Components of electromagnetic waves

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 $\boldsymbol{E}$ and the magnetic field component $\boldsymbol{B}$ are normal to the direction of propagation $\boldsymbol{v}$ of the wave:

$$\boldsymbol{\nu} = \boldsymbol{E} \times \boldsymbol{B} / |\boldsymbol{E} \times \boldsymbol{B}|$$

- $\boldsymbol{E}$ is normal to $\boldsymbol{B}$:

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

- The wave travels at the speed of light in a vacuum
- The fields $\boldsymbol{E}$ and $\boldsymbol{B}$ 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 λ 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.

Chapter 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 ν 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

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

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
>
> $$E = h\nu \quad , \qquad p = h/\lambda \text{ "}$$

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 $h\nu$, ν 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 ν 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 ν, 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 ν' of the scattered wave is smaller than the frequency ν 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 θ 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

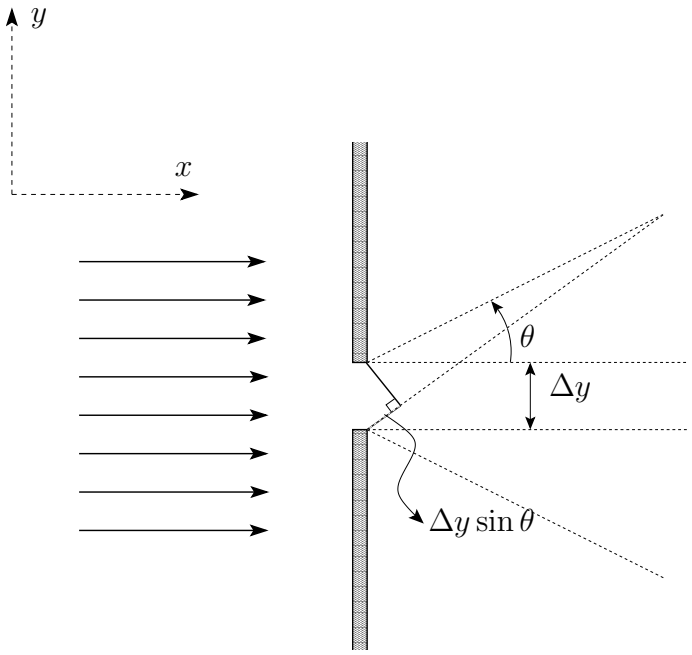


Figure 2.1: Diffraction of a beam of electrons

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 Δy, the beam is accompanied by diffraction and diverges by an angle θ. The electron is represented by a de Broglie wave of wave length $\lambda = h/p$. Thus,

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

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

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

Thus,

$$\boxed{\Delta y \Delta p_y \approx h} \tag{2.1}$$

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 Δp_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

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

$$\Delta y \Delta p \geq \frac{1}{2} h$$

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

2. $$\Delta t \Delta E \geq h$$

   Thus, just as position and momentum cannot be localized in time, a change in energy ΔE 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 Ψ. The central goal of quantum theory, then, is this: knowing Ψ at an initial time t_0, determine Ψ at later times t. To accomplish this, we must derive an equation of propagation of the wave Ψ, 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 began 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 ψ_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 ω 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,

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

Thus, E and p can be viewed as operators on Ψ:

$$\left.\begin{array}{rcl} E\Psi & = & \left(-\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial t}\right)\Psi \\ p\Psi & = & \left(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\right)\Psi \end{array}\right\} \tag{2.2}$$

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:

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

This is *Schrödinger's equation* for a free particle.

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

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

to obtain

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

Δ = Laplacian. What is ϕ? Let $t = 0$. Then the initial condition is

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

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

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

and Schrödinger's equation becomes,

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

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

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

and Schrödinger's equation becomes

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

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 $\Psi = \Psi(\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_Z, t)$.

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

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

From this we deduce the equation

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

This is called the Klein-Gordan equation. Henceforth, we restrict ourselves to non-relavistic quantum mechanics.

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,

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

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

$$E = \mathrm{i}\hbar\frac{\partial}{\partial t} \qquad \text{and} \qquad \boldsymbol{p}_r = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial \boldsymbol{q}_r} \;\; ; \qquad r = 1, 2, \ldots, N \tag{2.9}$$

Schrödenger's equation becomes,

$$\boxed{\left\{\mathrm{i}\hbar\frac{\partial}{\partial t} - H\left(\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_N; \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial \boldsymbol{q}_1}, \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial \boldsymbol{q}_2}, \ldots, \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial \boldsymbol{q}_N}\right)\right\}\Psi(\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_N, t) = 0} \tag{2.10}$$

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. The major source here is Griffith's book, *Introduction to Quantum Mechanics* [6], but we also consult the books of Born [2] and Messiah [13].

---

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

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

where

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

To (2.11) we must add an initial condition

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

where φ is prescribed. The wave function has the property

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

where $\Psi^* =$ the complex conjugate of Ψ and

$$\begin{aligned} \rho(x,t) &= \text{ the probability distribution function} \\ &\quad \text{associated with } \Psi, \text{ such that} \\ \rho(x,t)\mathrm{d}x &= \text{ the probability of finding the particle} \\ &\quad \text{between } x \text{ and } x + \mathrm{d}x \text{ at time } t \end{aligned}$$

The wave function must be normalized since ρ is a PDF:

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

Postulate. 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.

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

But

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

So

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

Hence,

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

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,

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

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

$$\begin{aligned}
\frac{\mathrm{d}\langle x \rangle}{\mathrm{d}t} &= \int_{-\infty}^{\infty} x \frac{\partial}{\partial t}(\Psi^* \Psi) \mathrm{d}t \\
&= \frac{\mathrm{i}\hbar}{2m} \int_{-\infty}^{\infty} x \frac{\partial}{\partial x} \left(\Psi^* \frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x} \Psi \right) \mathrm{d}x \qquad \text{(from (2.15))} \\
&= -\frac{\mathrm{i}\hbar}{2m} \int_{-\infty}^{\infty} \left(\Psi^* \frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x} \Psi \right) \mathrm{d}x \qquad \text{(integrating by parts)} \\
&= -\frac{\mathrm{i}\hbar}{m} \int_{-\infty}^{\infty} \Psi^* \frac{\partial \Psi}{\partial x} \mathrm{d}x \\
&= \frac{1}{m} \int_{-\infty}^{\infty} \Psi^* \left(\frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial x} \right) \Psi \mathrm{d}x \\
&= \frac{1}{m} \int_{-\infty}^{\infty} \Psi^* \rho \Psi \mathrm{d}x
\end{aligned}$$

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

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

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

The Heisenberg uncertainty principle can now be stated as,

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

where

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

2.9 Wave Packets / Fourier Transforms

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

$$\begin{aligned}
\Psi(x,t) &= \psi(x) \mathrm{e}^{\mathrm{i}\omega t} \\
&= \psi(x) \mathrm{e}^{-\mathrm{i}Et/\hbar}
\end{aligned} \tag{2.19}$$

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

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

where

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

This particular solution is not normalizable ($\int_{\mathbb{R}} \Psi^*\Psi \mathrm{d}x \to \infty$), 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,

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

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

$\Longrightarrow$

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

Equation (2.23) characterizes the wave as a sum of *wave packets*. It is a sinusoidal function modulated by the function φ. The wave function is the Fourier transform of φ and φ is the inverse Fourier transform of the initial value $\Psi(x,0)$ of Ψ. 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 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 $(\boldsymbol{p}, \boldsymbol{d}+\mathrm{d}\boldsymbol{p})$, we consider, for fixed time t, the Fourier transform $\Phi(\boldsymbol{p})$ of the wave function Ψ:

$$\left.\begin{array}{rcl} \Phi(\boldsymbol{p}) & = & \frac{1}{(2\pi\hbar)^{n/2}} \int_{\mathbb{R}^n} \Psi(\boldsymbol{x}) \mathrm{e}^{-\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{x}} \mathrm{d}x \\ \Psi(\boldsymbol{x}) & = & \frac{1}{(2\pi\hbar)^{n/2}} \int_{\mathbb{R}^n} \Phi(\boldsymbol{p}) \mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{x}} \mathrm{d}p \end{array}\right\} \tag{2.26}$$

(with $\mathrm{d}x = \mathrm{d}x_1 \cdots \mathrm{d}x_N$, $\mathrm{d}p = \mathrm{d}p_1 \cdots \mathrm{d}p_N$). Thus, the wave function can be viewed as a linear combination of waves $\exp(\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{x}/\hbar)$ 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 $(\boldsymbol{p}, \mathrm{d}\boldsymbol{p})$ is

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

and we must have

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

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,

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

The interpretation of the probability densities associated with Φ and Ψ 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 N of equivalent systems with the same wave function Ψ are considered. A position measurement on each of them gives the probability density $\Psi^*(\boldsymbol{x})\Psi(\boldsymbol{x})$ 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:

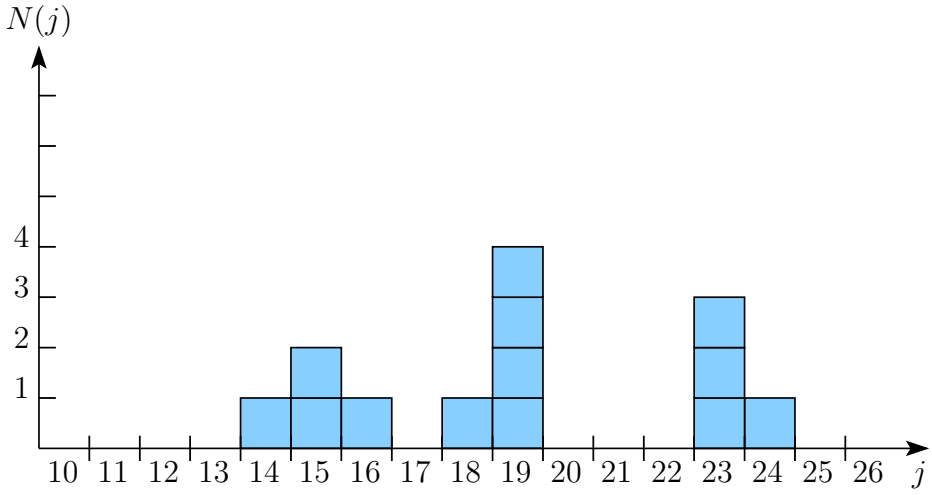


Figure 2.2: Histogram showing the number of people versus age.

1	aged	14
2	aged	15
1	aged	16
1	aged	18
4	aged	19
3	aged	23
1	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,

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

- The *mean* age, denoted $\langle j \rangle$:

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

Check it out:

$$\begin{aligned}
&\frac{(14) + 2(15) + (16) + (18) + 4(19) + 3(23) + (24)}{13} = \frac{247}{13} = 19 \\
&= 14\left(\frac{1}{13}\right) + 15\left(\frac{2}{13}\right) + 16\left(\frac{1}{13}\right) + 18\left(\frac{1}{13}\right) + 19\left(\frac{4}{13}\right) \\
&\quad + 23\left(\frac{3}{13}\right) + 24\left(\frac{1}{13}\right) \\
&= 14P(14) + 15P(15) + 16P(16) + 18P(18) + 19P(19) \\
&\quad + 23P(23) + 24P(24) \\
&= \sum_{j=0}^{\infty} jP(j)
\end{aligned}$$

We easily verify that

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

In general

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

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

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

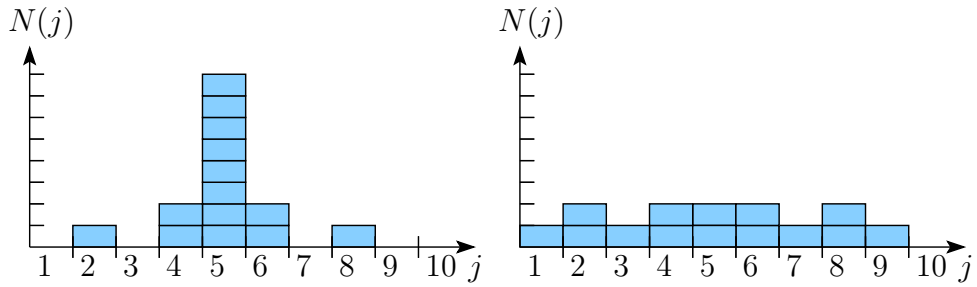


Figure 2.3: Two histograms with the same median, average, and probable value, but different standard deviations.

- It is easy to show that

$$\boxed{\sigma^2 = \langle j^2 \rangle - \langle j \rangle^2}$$

For the continuous case,

$$P_{ab} = \int_a^b \rho(x)\mathrm{d}x = \text{probability that } x \in [a,b]$$

$$\int_{-\infty}^{\infty} \rho(x)\mathrm{d}x = 1$$

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x)\rho(x)\mathrm{d}x$$

$$\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 \ , \qquad \langle x \rangle = \int_{-\infty}^{\infty} x\rho(x)\mathrm{d}x$$

Chapter 3

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 Ψ (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 powerful 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 x with 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 $\widetilde{p} = -i\hbar\frac{d}{dq}$). Thus, a function $F = F(\mathbf{q}, \mathbf{p})$ is associated with an operator $\widetilde{F}(\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N, -ih\frac{\partial}{\partial \mathbf{q}_1}, -i\hbar\frac{\partial}{\partial \mathbf{q}_2}, ..., -i\hbar\frac{\partial}{\partial \mathbf{q}_n})$ 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 $\mathbb{R}$, (Ψ 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 ψ^* is the complex conjugate of ψ. Thus $L^2(\mathbb{R})$ is a Hilbert space. The associated norm on $L^2(\mathbb{R})$ is then

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

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 $\mathbf{q}$ (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 as with the operators $\mathbf{p}_j = -i\hbar\partial/\partial\mathbf{q}_j$, dynamical variables likewise characterize operators,

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

(Analogously, one could define an operator

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

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

$$\begin{aligned} \langle Q \rangle &= \int_{\mathbb{R}} \psi^* Q \psi dq / \int_{\mathbb{R}} \psi^* \psi dq \\ &= \frac{\left\langle \Psi, \tilde{Q}\Psi \right\rangle}{\langle \Psi, \Psi \rangle} \end{aligned} \tag{3.1}$$

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

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

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.,

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

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

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

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

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

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

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

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

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

This is recognized as the *eigenvalue problem for the Hermitian operator* $\tilde{Q}$, with eigenvalue - eigenfunction pair (λ, ψ). 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 $\left\langle \psi, \tilde{Q}\psi \right\rangle = \langle Q \rangle = \lambda \langle \psi, \psi \rangle = \lambda$; 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 ψ 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 λ_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.*

$$\tilde{Q}\varphi_k = \lambda_k \varphi_k, \quad k = 1, 2, ... \tag{3.6}$$

Then

1. *if φ_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 λ_k, then any linear combination of these eigenfunctions is an eigenfunction corresponding to λ_k;*

3. *the eigenvalues are real*;

4. *the eigenfunctions φ_k can be used to construct an orthonormal set, i.e., a set of eigenfunctions of $\tilde{Q}$ such that*

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

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*

$$\lim_{m \to \infty} \parallel \psi - \sum_{k=1}^{m} c_k \varphi_k \parallel = 0 \tag{3.10}$$

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 (6) by φ_m and obtain the number

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

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

4. Consider,

$$\tilde{Q}\varphi_k = \lambda_k \psi_k \text{ and } \psi_m \varphi m \ \ , m \neq k$$

Then,

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

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

5. The eigenfunctions φ_k are assumed to be normalized:

$$\langle \varphi_k, \varphi_k \rangle = 1 \quad , k = 1, 2, ...$$

Thus they form an orthonormal basis for $L^2(\mathbb{R})$. Equation (3.8) is then just the Fourier representation of ψ 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). $\square$

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

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

and, in general

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

and, symbolically,

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

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

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

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,

$$\begin{aligned} e^{i\xi\tilde{Q}}\psi &= \sum_k c_k e^{i\xi\tilde{Q}}\varphi_k \\ &= \sum_k \sum_m c_k e^{i\xi\lambda m}\varphi_m\varphi_k \end{aligned}$$

and $\left|c_k e^{i\xi\lambda m}\right|^2 = |c_k|^2 \to 0$ 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 ϱ 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 $\varrho_1, \varrho_1, \cdots$ 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 λ_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 Ψ representing the dynamical state of the system,

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

where now

$$\begin{aligned} \varrho_k &= |c_k|^2 / \langle \Psi, \Psi \rangle \\ &= |\langle \varphi_k, \Psi \rangle|^2 / \langle \Psi, \Psi \rangle \end{aligned} \tag{3.16}$$

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

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*

$$\lambda_1, \lambda_2, \cdots \lambda_m, \cdots, \text{ and}$$

2. *the probability that Q takes on the value λ_k is*

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

where φ_k is the eigenfunction corresponding to λ_k and Ψ is the normalized wave function, $\langle \Psi, \Psi \rangle = 1$. □

We observe that the discrete probabilities ϱ_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)

$$pU(p^{'}, q) = p^{'}U(p^{'}, q) \tag{3.17}$$

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

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

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

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

In analogy with (3.9),

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

The analogue arguments cannot be carried further. Indeed, the function (19) 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 Ψ is

$$\langle Q \rangle = \Big\langle \Psi, \tilde{Q}\Psi \Big\rangle$$

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

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

For any other such operator M,

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

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

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

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 $[Q, M] \neq 0$), 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 ψ,

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

Hence,

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

or

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

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,

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

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

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

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

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

in analogy with (3.24).

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

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

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

Returning to the Hamiltonian operator H, if

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

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

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

Chapter 4

Selected Topics and Applications

4.1 Introductory Remarks

In this chapter, we collect discussions of special topics that complete an introduction to 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. The calculation of the quantum dynamics (the wave function) for the hydrogen atom is also described, and is a fundamental application of the theory developed thus far. We also consider the characterization of angular momentum and spin of particles in quantum systems.

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

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

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

The Hamiltonian, we recall, is of the form

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

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 ω:

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

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 power-series expansions (method of Frobenius) to be of the form

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

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

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

Thus, energy occurs in *quantized states* distinguished by different integer values of n, and the lowest energy is the ground state corresponding to $n = 0$.

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 the concrete example of the wave function for the hydrogen atom and the potential is that characterizing 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:

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

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

$$\triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{4.12}$$

Setting

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

we arrive at the time-independent equation,

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

The general solution is then

$$\Psi(x,y,z,t) = \sum_{n=1}^{\infty} c_n \psi_n(\mathbf{x}) e^{-iE_n t/\hbar} \tag{4.15}$$

with $\mathbf{x} = (x,y,z)$

In spherical coordinates (r,θ,ϕ) the time-independent equation (4.14) becomes

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2 sin\theta}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2 sin^2\theta}\left(\frac{\partial^2\psi}{\partial\phi^2}\right)\right] + V\psi = E\psi \tag{4.16}$$

Using the method of separation of variables, we set

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

where $R(r)$ and $Y(\theta,\phi)$ satisfy the equations,

$$\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) - \frac{2mr^2}{\hbar^2}[V(r) - E]R = \ell(1+\ell)R \tag{4.18}$$

$$(sin\theta)^{-1}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial Y}{\partial\theta}\right) + (sin\theta)^{-2}\frac{\partial^2 Y}{\partial\phi^2} = -\ell(1+\ell)Y \tag{4.19}$$

where $\ell(1+\ell)$ is the separation constant (Cf. Griffiths, p.124); ℓ is a complex number that will be determined to be an integer.

The solution to the angular equation is

$$Y_\ell^m(\theta,\phi) = \alpha C_{\ell m} P_\ell^m(cos\theta) e^{im\phi} \tag{4.20}$$

where

$$\alpha = \begin{cases} (-1)^m & \text{for } m \geq 0 \\ 1 & \text{for } m < 0 \end{cases}$$

$$C_{\ell m} = \left\{\frac{(\ell - |m|)!}{(\ell + |m|)!}\cdot\frac{(1+2\ell)}{4\pi}\right\}^{1/2} \tag{4.21}$$

and $P_\ell^m(\cdot)$ are the Legrindre polynomials of degree (m,ℓ).
One has

$$\int_0^{2\pi}\int_0^{\pi} Y_\ell^m(\theta,\phi)^* Y_{\ell'}^{m'}(\theta,\phi) sin\theta d\theta d\phi = \delta_{mm'}\delta_{\ell\ell'} \tag{4.22}$$

The radial equation (4.18) cannot be solved until a potential V is specified. For the hydrogen atom, we may consider a motionless proton, the origin of our spherical coordinate system, and a much lighter electron of charge $-e$ in an orbit (shell) around the protons. For V we take the Coulomb potential

$$V(r) = -\frac{e^2}{4\pi\varepsilon_0} \cdot \frac{1}{r} \tag{4.23}$$

The radial equation then reduces to

$$-\frac{\hbar^2}{2m}u''(r) + \left[-\frac{e^2}{4\pi\varepsilon_0} \cdot \frac{1}{r} + \frac{\hbar^2}{2m}\frac{\ell(1+\ell)}{r^2}\right] u = Eu \tag{4.24}$$

where $u(r) = rR(r), u''(r) = d^2u(r)/dr^2$. One can show that the solution to this equation is of the form,

$$u(r) = r(R)r_{=}\varrho^{1+\ell}e^{-\varrho}\sum_{j=0}^{\infty} a_j\varrho^j \tag{4.25}$$

where

$$\begin{array}{l} \varrho = kr,\ k = [-2mE/\hbar]^{1/2}\ \ (E<0),\ \text{and} \\ a_{j+1} = a_j[2(j+\ell+1) - \varrho_0]/(j+1)(j+2\ell+2) \\ \varrho_0 = me^2/2\pi\varepsilon_0\hbar^2k \end{array} \tag{4.26}$$

The series (4.25) blows up unless the a_j terminates at some $j = j^*$. From the condition $a_{j^*+1} = 0$, we conclude that $2(j^* + \ell + 1) - \varrho_0 = 0$, or that $\varrho_0 = 2n$, where $n = 1 + \ell + j^* \in \mathbb{Z}$. The energy is

$$E = -\frac{\hbar^2k^2}{2m} = -\frac{me^4}{8\pi^2\varepsilon_0^2\hbar^2\varrho_0^2}$$

so the allowed energies are

$$\begin{aligned} E_n &= -\left[\frac{m}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\right]\frac{1}{n^2} \\ &= \frac{E_1}{n^2}\ ,\quad n = 1, 2, 3, \cdots \end{aligned} \tag{4.27}$$

and E_1 is the energy of the *ground state* of the electron.

The above analysis is reproduced from Griffiths, p.136, 137. Formula (4.28) is called the *Bohr formula* and characterizes the energy at various quantum states. In conclusion,

$$R_{n\ell}(r) = r^{-1}\varrho^{1+\ell}e^{-\varrho}v(\varrho) \tag{4.28}$$

$$\begin{aligned} v(\varrho) &= \text{ polynomial in } \varrho \text{ of degree } j^* = n-\ell-1 \\ &= \sum_{j=0}^{j^*} a_j \varrho^j \\ a_{j+1} &= 2(j+\ell+1-n)a_j/(1+j)(2+2\ell+j) \end{aligned} \tag{4.29}$$

Returning to (4.26),

$$\varrho_0 = \frac{m^2e^2}{2\pi\varepsilon_0\hbar^2 k} = 2(1+\ell+j^*) = 2n$$

so

$$k = \left(\frac{me^2}{4\pi\varepsilon_0\hbar^2}\right)\frac{1}{n} = \frac{1}{a_0 n}$$

where

$$a_0 = \frac{4\pi\varepsilon_0\hbar^2}{me^2} = 0.529 \text{ x } 10^{-10}m \tag{4.30}$$

The number a_0, in units of length, is called the *Bohr radius*.

In (4.27), the ground energy state E_1 can be calculated to be

$$E_1 = -\frac{m}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 = -13.6eV \tag{4.31}$$

Any perturbation of some stationary state of the hydrogen atom may cause a transition to some other stationary state. If energy is absorbed, such as might result from a photon hitting the atom, a higher energy state may be attained; or if energy is given off, a lower state may be attained. Such quantum jumps in energy are constantly occurring. The difference in energy levels between an initial state $n = n_i$ and a final state n_f is,
from (4.27),

$$E_i - E_f = -13.6eV\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right)$$

According to Planck's formula, the energy of a photon is

$$E_i - E_f = hv = h\frac{c}{\lambda}$$

so

$$\frac{1}{\lambda} = R\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \tag{4.32}$$

Here R is the *Rydberg constant*

$$\begin{aligned} R &= \frac{m}{4\pi c^2\hbar^3}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 \\ &= 1.097 \text{ x } 10^7 m^{-1} \end{aligned} \tag{4.33}$$

If the transition brings the atom to the ground state ($n_f = 1$), the wave length λ corresponds to the ultraviolet length for electromagnet waves; if $n_f = 2$, the visible light region is attained; $n_f = 3$ corresponds to infrared.

4.4 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,

$$\mathbf{L} = \mathbf{q} \times \mathbf{p} \qquad \text{and } \mathbf{S} = I\boldsymbol{\omega} \tag{4.34}$$

where I is an inertial quantity (such as the moment of inertia) and $\boldsymbol{\omega}$ is the spin velocity vector. In quantum mechanics, an electron has extrinsic angular momentum and an intrinsic property, its spin, which it carries in addition to $\mathbf{L}$.

Denoting

$$\tilde{p}_j = \frac{\hbar}{i}\frac{\partial}{\partial q_j} \quad , \; j = 1, 2, 3 \tag{4.35}$$

the components of the extrinsic angular momentum are characterized by the quantum operators,

$$L_j = \varepsilon_{rsj} q_r \frac{i}{\hbar}\frac{\partial}{\partial q_s} \tag{4.36}$$

where repeated indices are summed and ε_{rsj} is the permutation symbol.

One can easily show that

$$[L_1, L_2] = i\hbar L_3 \; , \; [L_2, L_3] = i\hbar L_1 \; , \; [L_3, L_1] = i\hbar L_2 \tag{4.37}$$

The operators L_i are therefore incompatible observables, and

$$\sigma^2_{L_1}\sigma^2_{L_2} \geq \left(\frac{1}{2i}\langle i\hbar L_3\rangle\right)^2 = \frac{\hbar^2}{4}\langle L_3\rangle^2 \quad , \text{ etc.} \tag{4.38}$$

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

$$L^2 = L_i L_i = L_1^2 + L_2^2 + L_3^2 \tag{4.39}$$

does commute and

$$[L^2, L_k] = 0 \quad , \; k = 1, 2, 3 \tag{4.40}$$

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

$$L^2\varphi = \lambda\varphi \quad \text{and} \quad L_1\varphi = \mu\varphi \tag{4.41}$$

The major properties of the eigenfunctions and eigenvalues in (4.41) are laid out in the following theorem:

Theorem 4.1. *The spherical harmonics* Y_ℓ^m (*recall* (4.20)) *are eigenfunctions of the extrinsic angular momentum operators* L^2 and L_1 (or L_2, L_3) for

$$\ell = 0;\; 1/2,\; 1,\; 3/2,\; \cdots$$
$$m = -\ell,\; -\ell+1,\; \cdots,\; \ell-1,\; \ell$$

Moreover,

$$L^2 Y_\ell^m = \hbar^2\ell(1+\ell)Y_\ell^m \tag{4.42}$$

and

$$L_1 Y_\ell^m = \hbar m Y_\ell^m \tag{4.43}$$

Proof (See[5], pgs. 146-149) $\square$

In the case of the spin $\mathbf{S}(=$ "$I\omega$"), we have similarly,

$$[S_1, S_2] = i\hbar S_3,\ [S_2, S_3] = i\hbar S_1,\ [S_3, S_1] = i\hbar S_2 \tag{4.44}$$

and

$$S^2 q_m^s = \hbar^2 s(s+1) q_m^s;\ S_1 q_m^s = \hbar m q_m^s \tag{4.45}$$

$$s = 0,\ 1/2,\ 1,\ 3/2, \cdots\ ;\ \ m = -s,\ -s+1, \cdots,\ s-1,\ s$$

The case $s = 1/2$ is canonical. There are these two eigenstates $q_{1/2}^{1/2}, q_{-1/2}^{1/2}$. Representing these as vectors $\mathbf{q}_+ = \binom{1}{0}$ (the spin up) and $\mathbf{q}_- = \binom{1}{0}$ (the spin down), every state is a linear combination, $a\mathbf{q}_+ + b\mathbf{q}_-$, and the spin operators $\mathbf{S}$ become 2 x 2 matrices:

$$S_1 = \frac{\hbar}{2}\boldsymbol{\sigma}_1\ ,\ S_2 = \frac{\hbar}{2}\boldsymbol{\sigma}_2\ ,\ S_3 = \frac{\hbar}{2}\boldsymbol{\sigma}_3 \tag{4.46}$$

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

$$\boldsymbol{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ \boldsymbol{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.47}$$

All of the S_i and S^2 are Hermitian operators. The eigenvectors

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\quad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.48}$$

are called *spinors*, and satisfy

$$\begin{aligned} S^2\chi_+ &= \frac{3}{4}\hbar^2\chi_+\ ,\ \ S^2\chi_- = \frac{3}{4}\hbar^2\chi_- \\ S_3\chi_+ &= \frac{\hbar}{2}\chi_+\ ,\ \ S_3\chi_- = \frac{\hbar}{2}\chi_- \end{aligned} \tag{4.49}$$

We close this section by noting that a spinning charged elementary particle creates a magnetic dipole with a *magnetic dipole moment* $\mathbf{m}$ proportional to the spin angular momentum:

$$\mathbf{m} = \mu\mathbf{S} \tag{4.50}$$

The constant μ is called the *gyromagnetic ratio*. When placed in a magnetic field $\mathbf{B}$, it experiences a torque $\mathbf{m}$ x $\mathbf{B}$ with energy $-\mathbf{m}\cdot\mathbf{B}$. The Hamiltonian of a spinning particle at rest is

$$H = -\mu \mathbf{B}\cdot\mathbf{S} \tag{4.51}$$

4.5 Variational Principle

We begin with a statement of a minimization principle.

Theorem 4.2. *Let H be the Hamiltonian of a quantum system and let E_g denote its ground-state energy. Then*

$$E_g \le \langle \psi, H\psi\rangle \quad \forall \psi \in L^2(\mathbb{R}^N), \quad \langle \psi,\psi\rangle = 1 \tag{4.52}$$

In other words,

$$E_g \le \langle H\rangle \tag{4.53}$$

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

$$\begin{aligned}
\langle H\rangle &= \langle \psi, H\psi\rangle \\
&= \left\langle \sum_k c_k\psi_k, H\sum_j c_j\psi_j \right\rangle \\
&= \sum_k\sum_j c_k^* c_j E_j \langle \psi_k,\psi_j\rangle \\
&= \sum_j |c_j|^2 E_j \\
&\ge E_g \sum_j |c_j|^2 = E_g \qquad \square
\end{aligned}$$

This variational principle provides a basis for 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

$$H = -\frac{\hbar^2}{2m}(\triangle_1 + \triangle_2) - \frac{e^2}{4\pi\varepsilon_0}\left(\frac{2}{r_1} + \frac{2}{r_2} - \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\right) \tag{4.54}$$

In the Ritz method, we introduce a trial function,

$$\psi_a(\mathbf{r}_1, \mathbf{r}_2) = \frac{a^{-3}}{\pi} e^{-(r_1+r_2)}/a \tag{4.55}$$

where a is a variational parameter. The function ψ_a is suggested by perturbation theory. The corresponding energy is

$$E(a) = \langle \psi_a, H\psi_a \rangle \tag{4.56}$$

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

$$E_a \approx -2\left(Z - \frac{5}{16}\right)^2 E_H \tag{4.57}$$

where $Z = 1/2\pi\varepsilon_0$ and E_H the ground energy of the hydrogen atom (Cf. Messiah, p.771). Numerically,

$$E_a = -76.6eV$$

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

Chapter 5

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

5.1 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.

5.2 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 $\mathbf{O}$ by a position vector $\mathbf{r}_i$, $i = 1, 2, \cdots, N$. The motion of each such particle is assumed to be governed by Newton's second law:

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = \mathbf{F}_i \quad , \quad i = 1, 2, \cdots, N \tag{5.1}$$

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

$$\mathbf{F}_i = -\frac{\partial V}{\partial \mathbf{r}_i} \quad , i = 1, 2, \cdots, N \tag{5.2}$$

so that finally

$$\boxed{m_i \frac{d^2\mathbf{r}_i}{dt^2} + \frac{\partial V}{\partial \mathbf{r}_i} = \mathbf{0} \quad , i = 1, 2, \cdots, N} \tag{5.3}$$

To (5.3) 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 Ω, these generally chosen to represent a periodic pattern of the motion of the system throughout $\mathbb{R}^3$. The solutions

$\{\mathbf{r}_i(t)\}_{i=1}^N$ (5.3) 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,

$$\begin{aligned} &H(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N\ ;\ \mathbf{p}_1,\mathbf{p}_2,\cdots,\mathbf{p}_N) \\ &=\sum_{i=1}^N \frac{1}{2m_i}\mathbf{p}_i\cdot\mathbf{p}_i + V(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N) \end{aligned} \tag{5.4}$$

($\mathbf{q}_i \equiv \mathbf{r}_i$). Then, instead of (5.3), we have the $2N$ first-order systems,

$$\boxed{\dot{\mathbf{r}}_i = \frac{\partial H}{\partial \mathbf{p}_i}\quad,\quad \dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{r}_i} \qquad i=1,2,\cdots,N} \tag{5.5}$$

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:

$$\begin{aligned} V(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N) = &\sum_{i=1}^N V_1(\mathbf{r}_i) + \sum_{\substack{i,j=1\\ j>i}}^N V_2(\mathbf{r}_i,\mathbf{r}_j) \\ &+ \sum_{\substack{i,j,k=1\\ j>i,k>i}}^3 V_3(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k) + \cdots \end{aligned} \tag{5.6}$$

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,

$$V(\mathbf{r}_i, \mathbf{r}_j) = 4\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right]$$
$$r_{ij} = |\mathbf{r}_i - \mathbf{r}_j| \tag{5.7}$$

σ being a constant that depends upon the atomic structure of the atom or molecule at positions $\mathbf{r}_i, \mathbf{r}_j$, and ε 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 $\partial V/\partial \mathbf{r}_i$ 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

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

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 $r = R$.

5.3 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 $\mathbf{r}_{ai}, \mathbf{r}_{bi}$, $i =$

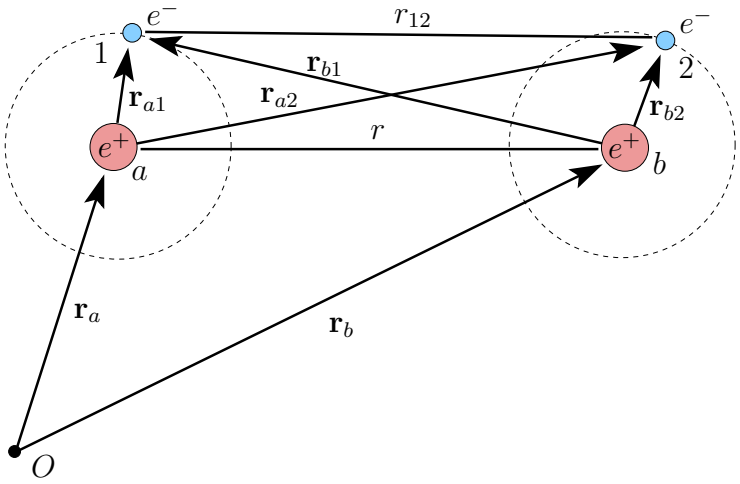


Figure 5.1: The hydrogen molecule H_2 consisting of two nuclei at a and b and electrons at 1 and 2.

$1, 2,$ and $\mathbf{r}_{12}$ is the distance between electrons. The distance separating the two atoms is denoted r. This situation is illustrated in Fig. 5.1.

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

$$\tilde{H} = -\frac{\hbar^2}{2m}\sum_{i=1}^{2}\triangle_i - \frac{1}{4\pi\varepsilon_0}\left\{\sum_{i=1}^{2}\left(\frac{e^2}{r_{ai}}+\frac{e^2}{r_{bi}}\right)+\frac{e^2}{r_{12}}+\frac{e^2}{r}\right\} \tag{5.8}$$

The corresponding Schrödinger equation is

$$\begin{aligned}\tilde{H}\Psi &= E\Psi \\ &= (V(r)+2E_1)\Psi\end{aligned} \tag{5.9}$$

Since

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

we have

$$V(r) = \langle\Psi, H\Psi\rangle - 2E_1 \tag{5.10}$$

where, we assume, the wave function Ψ depends parimetrically on r.

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

$$\boxed{\begin{aligned}\tilde{H}\Psi &= \left(V+\sum_{\alpha=1}^{M}E_\alpha\right)\Psi \\ m_i\ddot{\mathbf{r}}_i &= -\frac{\partial V}{\partial \mathbf{r}_i}\ ,\ i=1,2,\cdots,N\end{aligned}} \tag{5.11}$$

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 Ψ 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

$$E \approx \frac{\langle\varphi, H\varphi\rangle}{\langle\varphi,\varphi\rangle} = E_\varphi \tag{5.12}$$

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

$$V \approx E_{\varphi^*} - \sum_{\alpha=1}^{M} E_\alpha \tag{5.13}$$

We describe other approximate methods in the following sections.

5.4 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.

5.4.1 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 [7] and is attributed to Bloch [1] and Slater and Koster [9]. 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

$$H = -\frac{\hbar^2}{2m} \sum_{i=1}^{2} \triangle_i - \frac{1}{4\pi\varepsilon_o} \left\{ \sum_{\alpha=1}^{2} \left(\frac{e^2}{r_{\alpha 1}} + \frac{e^2}{r_{\alpha 2}} \right) - \frac{e^2}{r} \right\} \tag{5.14}$$

where we use the notation in Fig. 5.1. Let ψ_a and ψ_b be the molecular orbitals for a single hydrogen atom (i.e. $(-\hbar^2/2m\triangle - e^2/4\pi\varepsilon_0 r_{a1})\psi_a = 0$, $(-\hbar^2/2m\triangle - e^2/4\pi\varepsilon_0 r_{b2})\psi_b = 0$). Then, as an approximation of the wave function for H of (5.14), we take the linear combination,

$$\tilde{\Psi} = \alpha\psi_a + \beta\psi_b \tag{5.15}$$

The corresponding approximate energy is

$$\tilde{E} = \frac{\left\langle \tilde{\Psi}, H\tilde{\Psi} \right\rangle}{\left\langle \tilde{\Psi}, \tilde{\Psi} \right\rangle} = \frac{\alpha^2 H_{aa} + 2\alpha\beta H_{ab} + \beta^2 H_{bb}}{\alpha^2 + \beta^2 + 2\alpha\beta G_{\alpha\beta}} \tag{5.16}$$

where

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

Choosing α and β to minimize $\tilde{E}$, we have

$$\alpha =^{+}_{-} \beta$$

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

$$\psi^+ = A(\psi_a + \psi_b) \quad \text{and} \quad \psi^- = B(\psi_a - \psi_b) \tag{5.17}$$

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

$$\tilde{E}^+ = \frac{H_{aa} + H_{ab}}{1 + G_{ab}} \quad \text{and} \quad \tilde{E}^- = \frac{H_{aa} - H_{ab}}{1 - G_{ab}} \tag{5.18}$$

An approximation of the potential V can now be obtained from (5.13).

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.

5.4.2 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. 5.2. 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 $\mathbf{R}_k$, $k = 1, 2, \cdots, M$. Within the Bohr-Oppenheimer approximation, the problem of determining the ground state reduces to the nested minimization problems:

$$\inf_{\substack{\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_M \\ \in \; \mathbb{R}^{3M}}} \left\{ E(\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_M) + c \sum_{1 \le k < j \le M} \frac{Z_k Z_j}{R_{kj}} \right\} \tag{5.19}$$

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

$$E(\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_M) = \inf_{\psi \; \in \; \mathcal{H}} \{ \langle \psi, H\psi \rangle \; , \|\psi\| = 1 \} \tag{5.20}$$

where

$$H = -\sum_{i=1}^{N} \frac{\hbar^2}{2m} \triangle_{\mathbf{r}_i} - c \sum_{i=1}^{N} \left(\sum_{k=1}^{M} \frac{eZ_k}{|\mathbf{r}_j - \mathbf{R}_k|} \right) + c \sum_{1 \le i < j \le N} \frac{e^2}{r_{ij}} \tag{5.21}$$

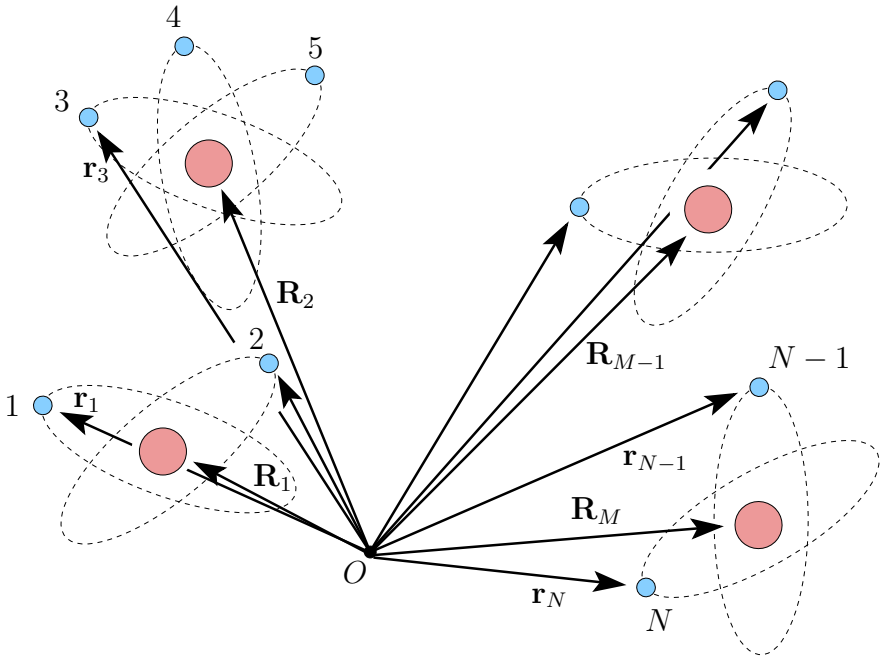


Figure 5.2: A molecular system consisting of M nuclei and N electrons.

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

$$\mathcal{H} = \prod_{i=1}^{N} H^1(\mathbb{R}^3, \mathbb{C}) \tag{5.22}$$

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 (5.20).

The Hartree-Fock approximation amounts to a variational formulation of (5.20) 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,

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

where φ_i are orthonormal molecular orbitals,

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

With this constraint, the Hartree-Fock energy functional becomes

$$\begin{aligned} E^{HF}(\{\varphi_i\}) &= \sum_{i=1}^{N} \frac{\hbar^2}{2m} \|\nabla_{\mathbf{r}_i} \varphi_i\|^2 + \int_{\mathbb{R}^3} \varrho_\varphi V dr \\ &+ \frac{c}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\varrho_\varphi(\mathbf{r}) \varrho_\varphi(\mathbf{r}^1)}{|\mathbf{r} - \mathbf{r}'|} dr dr' \\ &+ \frac{c}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \sum_{i=1}^{N} \varphi_i(\mathbf{r}) \varphi_i^*(\mathbf{r}) dr dr' \end{aligned} \tag{5.23}$$

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

$$\varrho_\varphi(\mathbf{r}) = \sum_{i=1}^{N} |\varphi_i(\mathbf{r})|^2$$

The Hartree-Fock reduced minimization thus reads,

$$\begin{aligned} \inf_{\varphi_i} \{ E^{HF}(\{\varphi_i\}) \ , \quad \varphi_i \in H^1(\mathbb{R}, \mathbb{C}), \ \langle \varphi_i, \varphi_j \rangle \\ = \delta_{ij} \quad , \quad i, j = 1, 2, \cdots, N \} \end{aligned} \tag{5.24}$$

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 (5.24). We construct a finite number of molecular orbits $\{\chi_m\}$, $1 \leq m \leq n$, (or approximate molecular orbits) and set

$$\varphi_i \approx \varphi_i^{(m)} = \sum_{k=1}^{m} C_{ki} \chi_k \tag{5.25}$$

where C_{ki} are constants to be determined so as to minimize E^{HF} over span $\{\varphi_i^{(m)},\ i = 1, 2, \cdots, N;\ m = 1, 2, \cdots, n\}$. 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 (5.25) into (5.23) leads to a reduced HF energy functional of the constants C_{ki}:

$$E^{EF}(\{\varphi_i^{(m)}\}) = \tilde{E}^{EF}(C_{kj}) \tag{5.26}$$

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

$$\frac{\partial \tilde{E}^{EF}}{\partial C_{kj}}(\hat{C}_{kj}) = 0 \tag{5.27}$$

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).

5.4.3 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 $\varrho(\mathbf{r})$ 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 $\varrho(\mathbf{r})$ is defined as the number of electrons per unit volume, and is given in terms of the wave function by (recalling that now Ψ is invariant under permutation)

$$\varrho(\mathbf{r}_1) = N \int_{\mathbb{R}} |\Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N)|^2 \, dx_2 dx_3 \cdots dx_N \tag{5.28}$$

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

$$\int_{\mathbb{R}^3} \varrho(\mathbf{r}) dr = N \tag{5.29}$$

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 φ_i for each electron:

$$\varrho = \sum_{i=1}^{N} |\varphi_i|^2 \tag{5.30}$$

For this system, the Hamiltonian operator is of the form

$$\tilde{H} = \tilde{T} + V_{ne} + V_{ee} \tag{5.31}$$

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

$$\tilde{T} = \frac{\hbar^2}{2m} \sum_{i=1}^{N} \triangle_{\mathbf{r}_i} \tag{5.32}$$

V_{ne} is the external potential for electron i due to nuclei α,

$$V_{ne} = -\sum_{i=1}^{N} \sum_{\alpha=1}^{M} \frac{cZ_\alpha e}{|\mathbf{r}_i - \mathbf{R}_\alpha|} \tag{5.33}$$

and

$$V_{ee} = \sum_{\substack{i,j=1 \\ i<j}}^{N} \frac{ce^2}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{5.34}$$

with c the usual constraint.

The electronic energy E satisfies, as usual,

$$\tilde{H}\Psi = E\Psi \tag{5.35}$$

and the total energy is

$$W = E + c \sum_{\substack{\alpha,\beta=1 \\ \alpha<\beta}}^{M} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} \tag{5.36}$$

$R_{\alpha\beta} = |\mathbf{R}_\alpha - \mathbf{R}_\beta|$.
Our goal is to determine these various energy components in terms of the density ϱ.

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

$$E(\varrho) = T(\varrho) + \frac{1}{2}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{\varrho(\mathbf{x})\varrho(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}dxdy + E_x(\varrho) + E_c(\varrho) \tag{5.37}$$

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

$$T(\varrho) = inf\left\{\frac{\hbar^2}{2m}\sum_{i=1}^{N}\int_{\mathbb{R}^3}|\nabla\varphi_i|^2\,dx\ ,\quad \varrho = \sum_{i=1}^{N}|\varphi_i|^2\right\} \tag{5.38}$$

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,

$$E_x(\varrho) = -c_x\int_{\mathbb{R}^3}\varrho^{4/3}dr \tag{5.39}$$

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 ϱ:

$$inf\left\{E(\varrho) + \int_{\mathbb{R}^3}\varrho V_{nec}dr\ ,\ \ \varrho \geq 0\ ,\ \int_{\mathbb{R}^3}\varrho dr = N\right\} \tag{5.40}$$

where V_{nec} is the electrostatic potential given by (5.6).

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

$$\begin{aligned} E^{TF}(\varrho) &= c_o\int_{\mathbb{R}^3}|\nabla\sqrt{\varrho}|^2\,dr + c_1\int_{\mathbb{R}^3}\varrho^{5/3}dr \\ &+ \frac{1}{2}\int_{\mathbb{R}^3\ \mathrm{x}\ \mathbb{R}^3}\frac{\varrho(\mathbf{x})\varrho(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}dxdy \\ &- \sum_{k=1}^{M}\int_{\mathbb{R}^3}\frac{\varrho(\mathbf{r})}{|\mathbf{r}-\mathbf{R}_k|}dr + \frac{1}{2}\sum_{\mathbf{x}\neq\mathbf{y}}\frac{ce^2}{|\mathbf{x}-\mathbf{y}|} \end{aligned} \tag{5.41}$$

and we seek,

$$inf\left\{E^{TF}(\varrho):\ \varrho\geq 0,\ \sqrt{\varrho}\ \in\ H^1(\mathbb{R}^3),\ \int_{\mathbb{R}^3}\varrho dr = N\right\} \tag{5.42}$$

5.5 Appendix: Identical Particles and the Pauli Exclusion Principle

Consider the Hamiltonian of a system of two identical particles,

$$H((\mathbf{r}_1, \mathbf{p}_1), (\mathbf{r}_1, \mathbf{p}_2)) = \frac{\mathbf{p}_1 \cdot \mathbf{p}_1}{2m} + \frac{\mathbf{p}_2 \cdot \mathbf{p}_2}{2m} + V(|\mathbf{r}_1 - \mathbf{r}_2|)$$

In classical mechanics, if these particles are identical, properties of the system cannot change if the particles are commuted; i.e. the states of particles are interchanged. Thus, H must be invariant under such a permutation ($H(\mu_1, \mu_2) = H(\mu_2, \mu_1)$).

In quantum mechanics, the situation is more serious. The following observations are classical Cf Messiah [13], Griffiths [6] or particularly Hinchliffe [8]. Suppose for a minute that the two electrons do not repel each other and that each moves freely about a nucleus of charge Z. The Hamiltonian splits into the sum of two parts,

$$(\hbar(\mathbf{r}_1) + \hbar(\mathbf{r}_2))\Psi(\mathbf{r}_1, \mathbf{r}_2) = E\Psi(\mathbf{r}_1, \mathbf{r}_2)$$

where

$$\hbar(\mathbf{r}_i) = -\frac{\hbar^2}{2m}\triangle_i - \frac{Ze^2}{4\pi\varepsilon_o}\,|\mathbf{R} - \mathbf{r}_i|$$

Using the method of separation of variables to solve this system, we find that the system breaks down to independent *orbitals*:

$$\hbar(\mathbf{r}_i)\varphi(\mathbf{r}_i) = E_i\varphi(\mathbf{r}_i)\ , \quad i = 1, 2$$

with $\Psi = \varphi_1(\mathbf{r}_1)\varphi_2(\mathbf{r}_2)$. This result of course is incorrect because we have neglected the electron–electron interaction.

Let φ_1 and φ_2 represent the lowest energy orbitals. Then possible states are

$$\Psi_{11}(\mathbf{r}_1, \mathbf{r}_2) = \varphi_1(\mathbf{r}_1)\varphi_1(\mathbf{r}_2) \quad E = 2E_1$$

$$\left.\begin{array}{rcl} \Psi_{12}(\mathbf{r}_1, \mathbf{r}_2) & = & \varphi_1(\mathbf{r}_1)\varphi_2(\mathbf{r}_2) \\ \Psi_{21}(\mathbf{r}_1, \mathbf{r}_2) & = & \varphi_2(\mathbf{r}_1)\varphi_1(\mathbf{r}_2) \end{array}\right\} E = E_1 + E_2$$

But the latter case presents a problem: the electrons, being indistinguishable, must appear on equal footing in computing these observables, which is not the case as electron 1 is apparently different from Z as it appears is a first-energy orbit in Ψ_{12} but in a second in Ψ_{21}. However, the combination,

$$\Psi_{12} + \Psi_{21} \quad \text{and} \quad \Psi_{12} - \Psi_{21}$$

do not have this shortcoming, and are conceivable bases for building more complex descriptions of possible dynamical states.

A wave function will be referred to as symmetric or antisymmetric according to whether it remains the same or changes sign under a change in electron labels. Thus, the first sum above is symmetric whilst the second is antisymmetric. For multielectron systems, we classify wave functions as symmetric or antisymmetric according to whether or not a ...?

Let us suppose that the wave function at time t is a linear combination of a symmetric and an antisymmetric component:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, t) = \alpha\psi_{sym} + \beta\psi_{ant}$$

and the probability density of finding one particle at $\mathbf{r}'$ and another at $\mathbf{r}''$ is

$$|\Psi(\mathbf{r}', \mathbf{r}'')|^2 + |\Psi(\mathbf{r}'', \mathbf{r}')|^2 = 2\,|\alpha|^2\,|\psi_{ant}|^2 + 2\,|\beta|^2\,|\psi_{sym}|^2$$

This expression must be independent of α and β (which are arbitrary constants). Thus, we must have,

$$|\psi_{ant}(\mathbf{r}', \mathbf{r}'')| = |\psi_{sym}(\mathbf{r}', \mathbf{r}'')|$$

This result leads us to the following postulate: *the dynamical states of two identical particles are either all symmetrical* ($\alpha = 1, \beta = 0$) *or all antisymmetrical* ($\alpha = 0, \beta = 1$) i*n the permutation of the two particles* (Cf. Messiah [8, p.585].

To extend this postulate to systems of N identical particles requires an introduction of permutations in N variables. In any case, the states of systems containing N particles are either all symmetrical or all antisymmetrical with respect to the permutations of N particles. Particles with symmetrical states are called *bosons* while those with antisymmetrical states are called *fermions*. According to Messiah [8, p.595], elementary particles with spin

1/2 occurring in nature (e.g. electrons, protons) are fermions while those with integral spin are bosons (e.g. photons). From what has been said, it can be shown to follow that *two identical fermions cannot occupy the same state*. This is the *Pauli Exclusion Principle*.

A method for constructing antisymmetrical wave functions is to use the *Slater determinant*. Let $\psi_1, \psi_2, \cdots, \psi_N$ denote N orthogonal individual states occupied by N particles. The corresponding antisymmetrical state is represented by the N x N Slater determinant,

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) \cdots \psi_N(\mathbf{r}_1) \\ \psi_2(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) \cdots \psi_N(\mathbf{r}_2) \\ \vdots & \vdots \\ \psi_N(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) \cdots \psi_N(\mathbf{r}_N) \end{vmatrix}$$

Bibliography

[1] Bloch, F. "Uber die quantenmechanik der elektron in kristallgittern," *Zeitschrift fur Physik*, 52, pp.522-600, 1928.

[2] Born, M. *Atomic Physics*, 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," *Mathematical Models and Methods for Ab Initio Quantum Chemistry*, 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.) *Mathematical Models and Methods in Ab Initio Quantum Chemistry*. Springer-Verlag, Berlin, Heidelberg, New York, 2000.

[5] Dreizler, R.M. and Gross, E.K.U. *Density Function Theory*, Springer-Verlag, Berlin, Heidelberg, New York, 1993.

[6] Griffiths, D.J. *Introduction to Quantum Mechanics*, Prentice-Hall, Upper Saddle River, NJ, 1995.

[7] Halliday, D., Resnick, R., and Walker, J. *Fundamentals of Physics*, 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. *Classical Electrodynamics*, 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," *Computer Methods in Applied Mechanics and Engineering*, Vol. 193, pp.1529-1578, 2004.

[13] Messiah, A. *Quantum Mechanics*, Volumes 1 and 2. Dover, Mineola, NY, 1999.

[14] Parr, R.G. and Yang, W. *Density Function Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.

[15] Slater, V.C., and Koster, G.F. "Simplified LCAO method for periodic potential problem," *Physics Review*, 94(6), pp.1498-1524, 1954.

[16] von Neumann, J. *Mathematical Foundations of Quantum Mechanics*, University Press, Princeton, NJ, 1955.