CHEM 409:   Quantum Chemistry II   **L2 P1T0   (2 Credits)**

Quantum chemistry including quantum mechanics and statistical thermodynamics, relationships between the various bonding theories-balance bond, molecular orbital, LCAO and non- paring spatial theory, postulates of the quantum theory. The Schrödinger equation and the particle in a box problem, treatment of the hydrogen molecule ion, and many electron atoms, perturbation and variation methods

Chapter One

1.1 Bond Theories

Valence Bond theory describes covalent bond formation as well as the electronic structure of molecules. The theory assumes that electrons occupy atomic orbitals of individual atoms within a molecule, and that the electrons of one atom are attracted to the nucleus of another atom. This attraction increases as the atoms approach one another until the atoms reach a minimum distance where the electron density begins to cause repulsion between the two atoms. This electron density at the minimum distance between the two atoms is where the lowest potential energy is acquired, and it can be considered to be what holds the two atoms together in a chemical bond. Valence bond (VB) theory assumes that all bonds are localized bonds formed between two atoms by the donation of an electron from each atom. This is actually an invalid assumption because many atoms bond using delocalized electrons

In chemistry, **valence bond (VB) theory** is one of two basic theories, along with molecular orbital (MO) theory, that were developed to use the methods of quantum mechanics to explain chemical bonding. It focuses on how the atomic orbitals of the dissociated atoms combine to give individual chemical bonds when a molecule is formed. In contrast, molecular orbital theory has orbitals that cover the whole molecule

1.2 LCAO

A **linear combination of atomic orbitals** or **LCAO** is a quantum superposition of atomic orbitals and a technique for calculating molecular orbitals in quantum chemistry. In quantum mechanics, electron configurations of atoms are described as wave functions. In mathematical sense, these wave functions are the basis set of functions, which describe the electrons of a given atom. In chemical reactions, orbital wave functions are modified, i.e. the electron cloud shape is changed, according to the type of atoms participating in the chemical bond.

It was introduced in 1929 by Sir John Lennard-Jones with the description of bonding in the diatomic molecules of the first main row of the periodic table, but had been used earlier by Linus Pauling for H_2^+.

A mathematical description follows.

An initial assumption is that the number of molecular orbitals is equal to the number of atomic orbitals included in the linear expansion. In a sense, n atomic orbitals combine to form n molecular orbitals, which can be numbered $i = 1$ to n and which may not all be the same. The expression (linear expansion) for the i th molecular orbital would be:

$$\phi_i = c_{1i}\chi_1 + c_{2i}\chi_2 + c_{3i}\chi_3 + \cdots + c_{ni}\chi_n$$

or

$$\phi_i = \sum_r c_{ri}\chi_r$$

where ϕ_i(phi) is a molecular orbital represented as the sum of n atomic orbitals χ_r(chi), each multiplied by a corresponding coefficient c_{ri}, and r (numbered 1 to n) represents which atomic orbital is combined in the term. The coefficients are the weights of the contributions of the n atomic orbitals to the molecular orbital. The Hartree–Fock procedure is used to obtain the coefficients of the expansion.

The orbitals are thus expressed as linear combinations of basis functions, and the basis functions are one-electron functions centered on nuclei of the component atoms of the molecule. The atomic orbitals used are typically those of hydrogen-like atoms since these are known analytically i.e. Slater-type orbitals but other choices are possible like Gaussian functions from standard basis sets.

By minimizing the total energy of the system, an appropriate set of coefficients of the linear combinations is determined. This quantitative approach is now known as the Hartree–Fock method. However, since the development of computational chemistry, the LCAO method often refers not to an actual optimization of the wave function but to a qualitative discussion which is very useful for predicting and rationalizing results obtained via more modern methods. In this case, the shape of the molecular orbitals and their respective energies are deduced approximately from comparing the energies of the atomic orbitals of the individual atoms (or molecular fragments) and applying some recipes known as level repulsion and the like. The graphs that are plotted to make this discussion clearer are called **correlation diagrams**. The required atomic orbital energies can come from calculations or directly from experiment via Koopmans' theorem.

1.3 Principles of Molecular Orbital Theory

In molecules, atomic orbitals combine to form molecular orbitals which surround the molecule. Similar to atomic orbitals, molecular orbitals are wave functions giving the probability of finding an electron in certain regions of a molecule. Each molecular orbital can only have 2 electrons, each with an opposite spin. The molecular orbitals and their energy ordering the ground state configuration is found by applying the Pauli principle, the Aufbau principle and Hund's rule just as with atoms.

The principles to apply when forming pictorial molecular orbitals from atomic orbitals are summarized in the table below:

Principle	Details/Examples
Total number of molecular orbitals is equal to the total number of atomic orbitals used to make them.	The molecule H_2 is composed of two H atoms. Both H atoms have a 1s orbital, so when bonded together, there are therefore two molecular orbitals.
Bonding molecular orbitals are lower energy than the atomic orbitals from which they were formed.	Electrons in bonding molecular orbitals help stabilize a system of atoms since less energy is associated with bonded atoms as opposed to a system of unbound atoms. Bonding orbitals are formed by in-phase combinations of atomic orbitals and increase the electron density between the atoms (see figure 2 below)
Antibonding molecular orbitals are higher energy than the atomic orbitals from which they were formed.	Electrons in antibonding molecular orbitals cause a system to be destabilized since more energy is associated with bonded atoms than that of a system of unbound atoms. Antibonding orbitals are formed by out-of-phase combinations of atomic orbitals and decrease the electron density between atoms (see figure 2 below).
Following both the Pauli exclusion principle and Hund's rule, electrons fill in orbitals of increasing energy.	Electrons fill orbitals with the lowest energy first. No more than 2 electrons can occupy 1 molecular orbital at a time. Furthermore, all orbitals at an energy level must be filled with one electron before they can be paired. (see figure 3 below)
Molecular orbitals are best formed when composed of Atomic orbitals of like energies .	When Li_2 forms the two lowest energy orbitals are the pair of bonding and antibonding orbitals formed from the two possible combinations of the 1s on each atom. The 2s orbitals combine primarily with each other to form another pair of bonding and antibonding orbitals at a higher energy.

Figure 2 (below) shows how bonding and antibonding σ orbitals can be formed by combining s orbitals in-phase (bonding, bottom) and out-of-phase (antibonding, top). If the atomic orbitals are combined with the same phase they interfere constructively and a bonding orbital is formed. Bonding molecular orbitals have lower energy than the atomic orbitals from which they were formed. The lowering of the energy is attributed to the increase in shielding of the nuclear repulsion because of the increase in electron density between the nuclei. If the atomic orbitals are combined with different phases, they interfere destructively and an antibonding molecular orbital is formed. Antibonding molecular orbitals have a higher energy than the atomic orbitals from which they were formed. The higher energy is attributed to the reduced shielding of the nuclear repulsion because of the lower electron probability density between the nuclei.
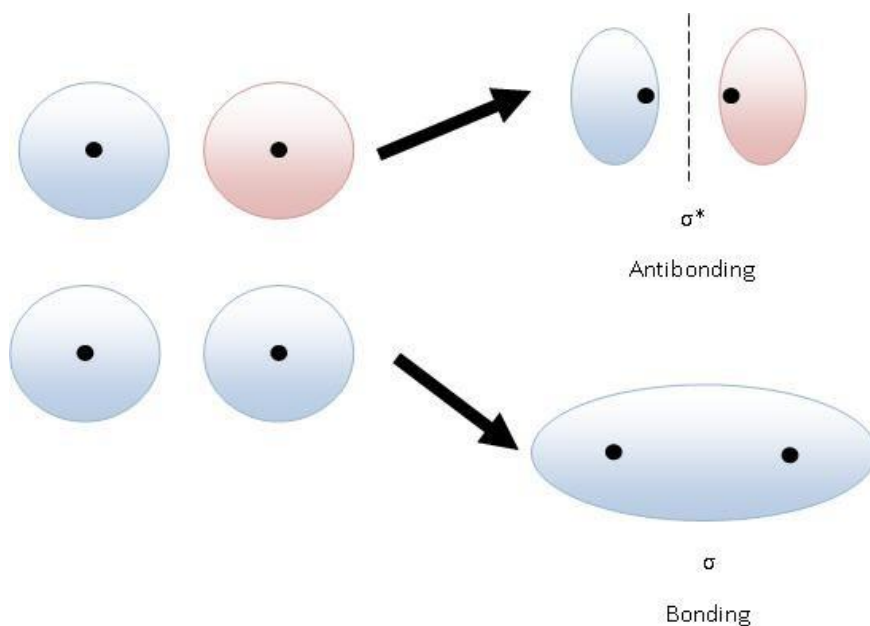
Figure 2: Combining hydrogen-like s orbitals to generate bonding (bottom) and antibonding (top) orbitals. The dark dot represents the location of the nucleus. Note the decrease in electron density between the nuclei in the antibonding orbital.

σ Bonds

Molecular orbitals that are symmetrical about the axis of the bond are called sigma molecular orbitals, often abbreviated by the Greek letter σ . Figure 2 shows the 1s orbitals of 2 Hydrogen atoms forming sigma orbitals. There are two types of sigma orbitals formed, antibonding sigma orbitals (abbreviated $\sigma *$), and bonding sigma orbitals (abbreviated σ). In sigma bonding orbitals, the in phase atomic orbitals overlap causing an increase in electron density along the bond axis. Where the atomic orbitals overlap, there is an increase in electron density and therefore an increase in the intensity of the negative charge. This increase in negative charge causes the nuclei to be drawn closer together. In sigma antibonding orbitals ($\sigma *$), the out of phase 1s orbitals interfere destructively which results in a low electron density between the nuclei as seen on the top of the diagram.

The diagram below (figure 3) is a representation of the energy levels of the bonding and antibonding orbitals formed in the hydrogen molecule. Two molecular orbitals were formed: one antibonding (σ^*) and one bonding (σ).The two electrons in the hydrogen molecule have antiparallel spins. Notice that the σ^* orbital is empty and has a higher energy than the σ orbital.
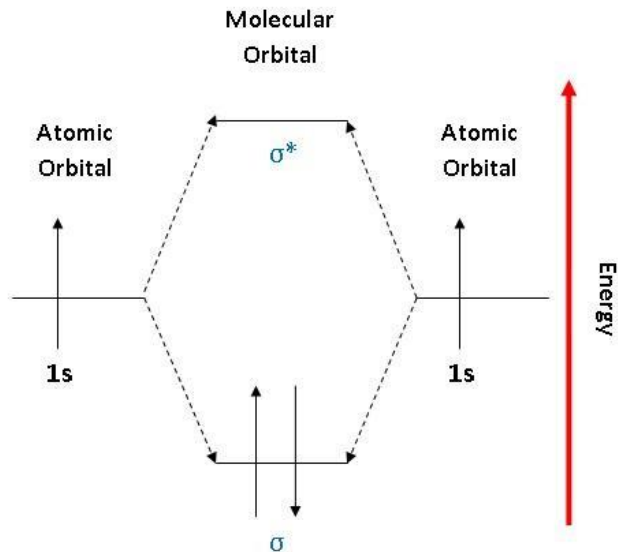
Figure 3: An MO energy level diagram for H_2. The up and down arrows represent electrons that are spin up or spin down.

Sigma bonding orbitals and antibonding orbitals can also be formed between p orbitals (figure 4). Notice that the orbitals have to be in phase in order to form bonding orbitals. Sigma molecular orbitals formed by p orbitals are often differentiated from other types of sigma orbitals by adding the subscript p below it. So the antibonding orbital shown in the diagram below would be σ^*_p.
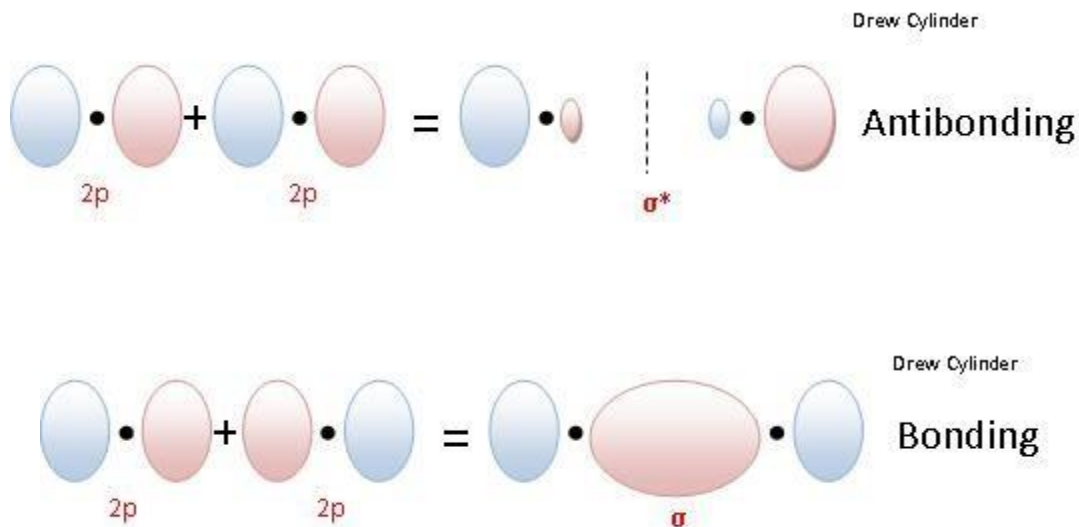

Figure 4: The formation of a σ bonding and antibonding orbital using p-orbitals.

π Bonds

The π bonding is a side to side overlap of orbitals, which then causes there to be no electron density along the axis, but there is density above and below the axis. The diagram below (figure 5) shows a π antibonding molecular orbital and a π bonding molecular orbital.
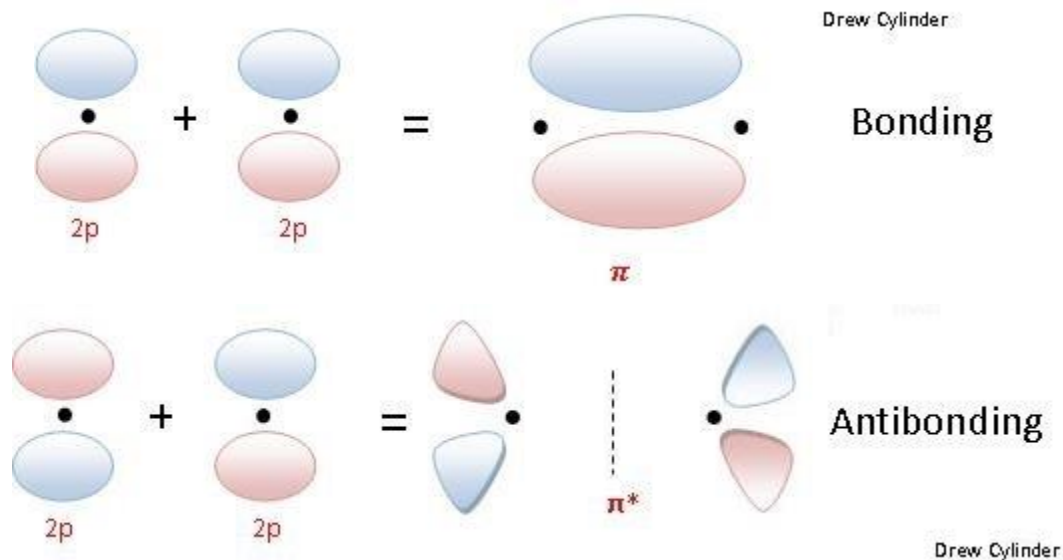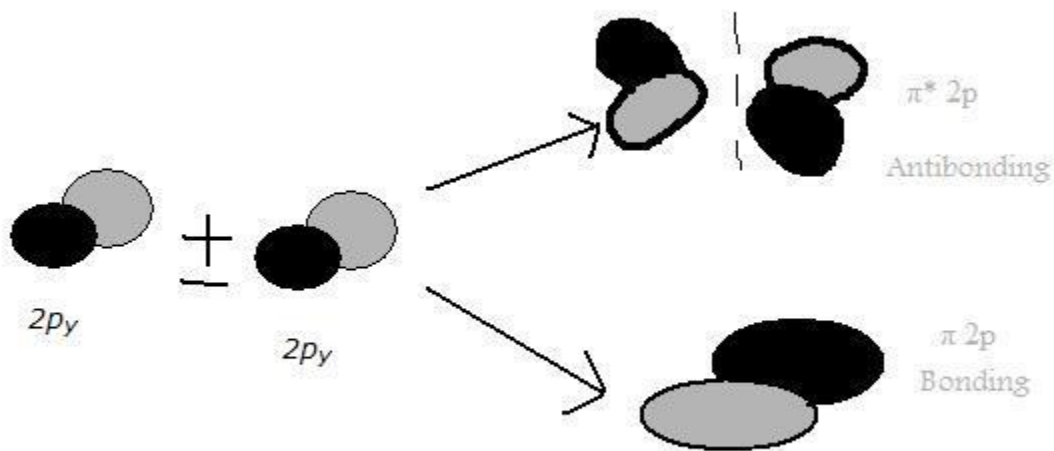

Figure 5: The side on overlap of p orbitals to form pi bonding and antibonding orbitals. Note that there is a second set of p orbitals sticking in and out of the image that can combine in the same way. (see cartoons immediately below)

2p$_y$ Orbitals

The two 2p$_y$ atomic orbitals overlap in parallel to form two π molecular orbitals which are asymmetrical about the axis of the bond.

2p$_z$ orbitals

π* 2p
Antibonding

2p$_z$ \quad 2p$_z$

π 2p
Bonding
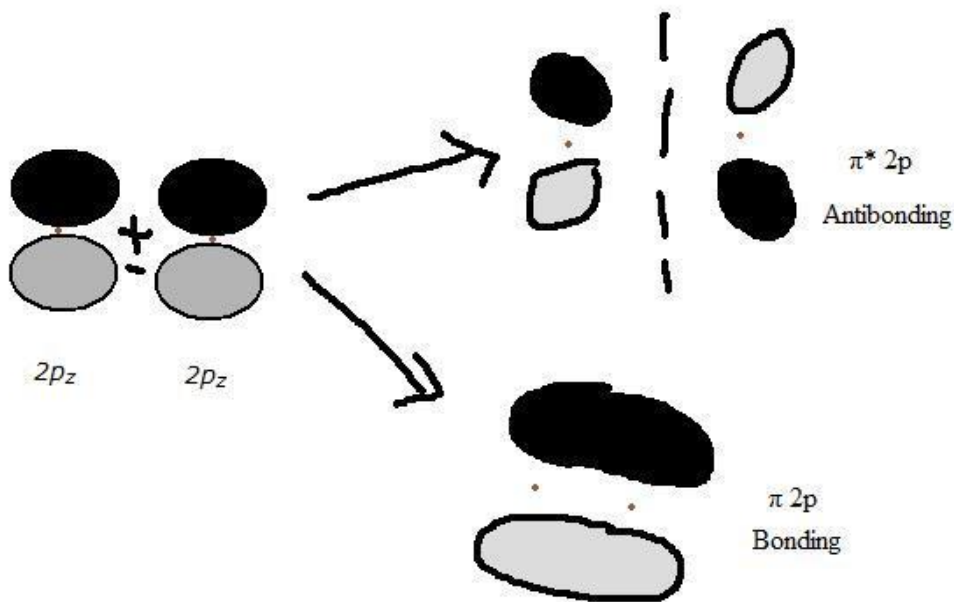
The two 2p$_z$ orbitals overlap to create another pair of pi 2p and pi *2p molecular orbitals. The 2p$_z$-2p$_z$ overlap is similar to the 2p$_y$-2p$_y$ overlap because it is just the orbitals of the 2pz rotated 90 degrees about the axis. The new molecular orbitals have the same potential energies as those from the 2p$_y$-2p$_y$ overlap.

In summary the three pairs of p orbitals can combine to form one set of \sigma\ orbitals and two sets of \pi\ orbitals.

Bond Orders and Stability of Molecules

Bond Order indicates the strength of the bond with the greater the bond order, the stronger the bond.

$$\text{Bond Order} = \frac{1}{2}(a-b)$$

where

- a is the number of electrons in bonding molecular orbitals and
- b is the number of electrons in antibondng molecular orbitals.

If the bond order is zero, then no bonds are produced and the molecule is not stable (for example He_2). If the Bond Order is 1, then it is a single covalent bond. The higher the Bond Order, the

more stable the molecule is. An advantage of Molecular Orbital Theory when it comes to Bond Order is that it can more accurately describe partial bonds (for example in H_2^+, where the Bond Order=1/2), than Lewis Structures

1.4 Comparison of Theories

Valence bond theory
1- In valence bond theory, two atomic orbitals give an inter-atomic orbital obtained by space filling two unpaired electrons, one being in each of the two atomic orbitals.
2- The resulting molecule, consist of atoms and retain their individual character.
3- Atomic orbitals are mono-centric.
Molecular orbital theory
1- In molecular orbital theory, molecular orbits are formed by linear combination of atomic orbitals) method.
2- Atomic orbitals of the resulting molecule lose their individual identities.
3- Molecular orbitals are poly-centric
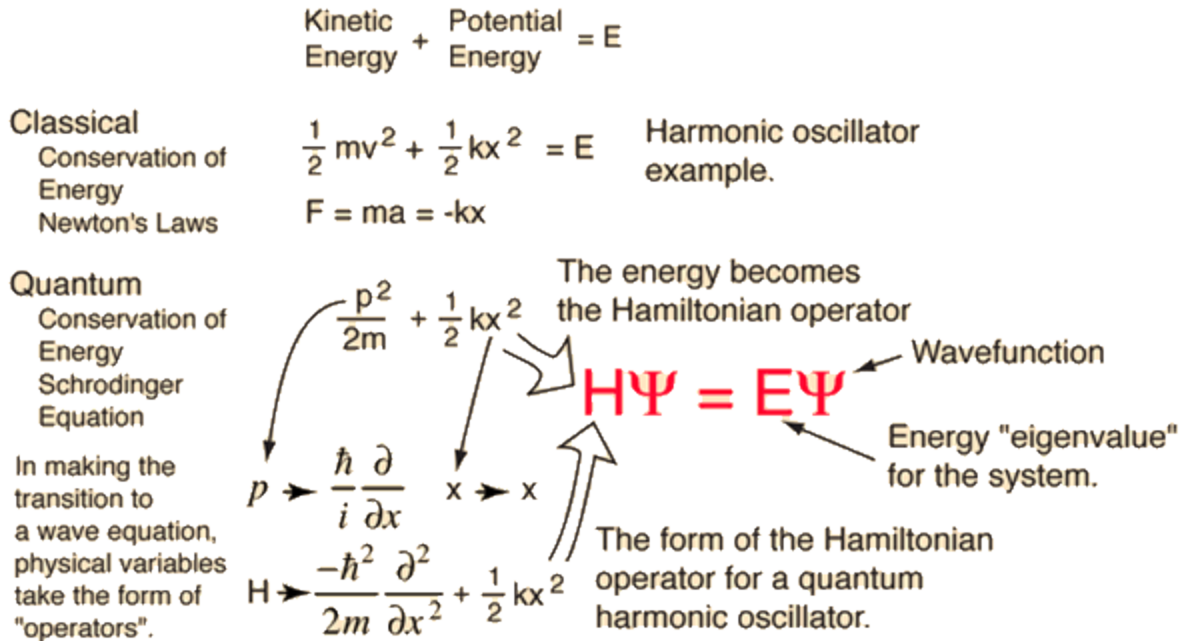
In LCAO, atomic orbitals are superimposed on each other

Chapter Two

2.1 Quantum Mechanics

Quantum mechanics is the theoretical framework which describes the behavior of matter on the atomic scale

2.2 Schrodinger Equation

The Schrodinger equation plays the role of Newton's laws and conservation of energy in classical mechanics - i.e., it predicts the future behavior of a dynamic system. It is a wave equation in terms of the wave function which predicts analytically and precisely the probability of events or outcome. The detailed outcome is not strictly determined, but given a large number of events, the Schrodinger equation will predict the distribution of results.

$$\frac{\text{Kinetic}}{\text{Energy}} + \frac{\text{Potential}}{\text{Energy}} = E$$

Classical
Conservation of Energy
Newton's Laws

$$\frac{1}{2}mv^2 + \frac{1}{2}kx^2 = E$$

$$F = ma = -kx$$

Harmonic oscillator example.

Quantum
Conservation of Energy
Schrodinger Equation

$$\frac{p^2}{2m} + \frac{1}{2}kx^2$$

The energy becomes the Hamiltonian operator

$$H\Psi = E\Psi$$

Wavefunction

Energy "eigenvalue" for the system.

In making the transition to a wave equation, physical variables take the form of "operators".

$$p \rightarrow \frac{\hbar}{i}\frac{\partial}{\partial x}$$

$$x \rightarrow x$$

$$H \rightarrow \frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}kx^2$$

The form of the Hamiltonian operator for a quantum harmonic oscillator.

The kinetic and potential energies are transformed into the Hamiltonian which acts upon the wave function to generate the evolution of the wave function in time and space. The Schrodinger equation gives the quantized energies of the system and gives the form of the wave function so that other properties may be calculated

2.3 Free particle approach to the Schrodinger equation

Though the Schrodinger equation cannot be derived, it can be shown to be consistent with experiment. The most valid test of a model is whether it faithfully describes the real world. The wave nature of the electron has been clearly shown in experiments like the Davisson-Germer experiment. This raises the question "What is the nature of the wave?". We reply, in retrospect, that the wave is the wave function for the electron. Starting with the expression for a traveling wave in one dimension, the connection can be made to the Schrodinger equation. This process makes use of the de Broglie ($\lambda= h/p$) relationship between wavelength and momentum and the Planck relationship between frequency and energy.

Ψ = wavefunction for electron

$$\Psi = A\cos\left(\frac{2\pi}{\lambda}x - \omega t\right)$$

Using the deBroglie relationship

$$\frac{2\pi}{\lambda} = \frac{2\pi p}{h} = k$$

p = electron momentum

Using the Planck relationship

$$\omega = \frac{\hbar\omega}{\hbar} = \frac{E}{\hbar}$$

E = electron energy

It is easier to show the relationship to the Schrodinger equation by generalizing this wave function to a complex exponential form using the Euler relationship. This is the standard form for the free particle wave function.

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

If we now take the partial derivatives of this wave function with respect to position and time, we can show that these derivatives are related to momentum and energy respectively.

$$\frac{\partial \Psi}{\partial x} = ik\Psi = \frac{ip}{\hbar}\Psi$$

The momentum connection

$$\frac{\partial \Psi}{\partial t} = -i\omega\Psi = \frac{-iE}{\hbar}\Psi$$

The energy connection

When an operation on a function gives back a constant times the function, that constant is called an eigenvalue, and the function is an eigen function. The above relationships can be rearranged as follows.

$$P_{operator}\Psi = -i\hbar\frac{\partial \Psi}{\partial x} = p\Psi$$

$$E_{operator}\Psi = i\hbar\frac{\partial \Psi}{\partial t} = E\Psi$$

$$P_{operator} = -i\hbar\frac{\partial}{\partial x}$$

$$E_{operator} = i\hbar\frac{\partial}{\partial t}$$

In this kind of fashion, quantum mechanical "operators" can be developed for relevant physical observables.

The connection to the Schrodinger equation can be made by examining wave and particle expressions for energy:

$$\frac{1}{2}mv^2 = \frac{p^2}{2m} \overset{particle}{\Longleftarrow} E \overset{wave}{\Longrightarrow} h\nu = \hbar\omega$$

Asserting the equivalence of these two expressions for energy and putting in the quantum mechanical operators for both brings us to the Schrodinger equation

$$\frac{-\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} = i\hbar\frac{\partial\Psi}{\partial t}$$

Schrodinger equation for free particle
in one dimension.

2.4 Particle in a box
Some trajectories of a particle in a box according to Newton's laws of classical mechanics (A), and according to the Schrödinger equation of quantum mechanics (B-F). In (B-F), the horizontal axis is position, and the vertical axis is the real part (blue) and imaginary part (red) of the wavefuntion. The states (B,C,D) are energy eigenstates, but (E,F) are not. In quantum mechanics, the **particle in a box** model (also known as the **infinite potential well** or the **infinite square well**) describes a particle free to move in a small space surrounded by impenetrable barriers. The model is mainly used as a hypothetical example to illustrate the differences between classical and quantum systems. In classical systems, for example a ball trapped inside a large box, the particle can move at any speed within the box and it is no more likely to be found at one position than another. However, when the well becomes very narrow (on the scale of a few nanometers), quantum effects become important. The particle may only occupy certain positive energy levels. Likewise, it can never have zero energy, meaning that the particle can never "sit still". Additionally, it is more likely to be found at certain positions than at others, depending on its energy level. The particle may never be detected at certain positions, known as spatial nodes.

The particle in a box model provides one of the very few problems in quantum mechanics which can be solved analytically, without approximations. This means that the observable properties of the particle (such as its energy and position) are related to the mass of the particle and the width of the well by simple mathematical expressions. Due to its simplicity, the model allows insight into quantum effects without the need for complicated mathematics. It is one of the first quantum mechanics problems taught in undergraduate physics courses, and it is commonly used as an approximation for more complicated quantum systems

2.5 One-dimensional solution
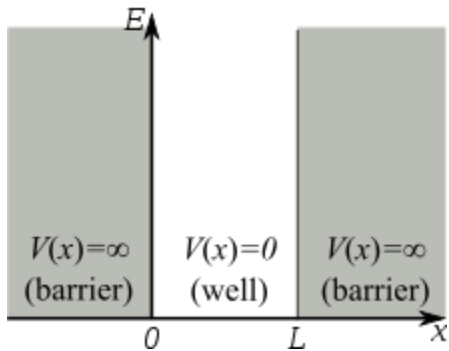
The barriers outside a one-dimensional box have infinitely large potential, while the interior of the box has a constant, zero potential.

The simplest form of the particle in a box model considers a one-dimensional system. Here, the particle may only move backwards and forwards along a straight line with impenetrable barriers at either end.[1] The walls of a one-dimensional box may be visualized as regions of space with an infinitely large potential energy. Conversely, the interior of the box has a constant, zero potential energy.[2] This means that no forces act upon the particle inside the box and it can move freely in that region. However, infinitely large forces repel the particle if it touches the walls of the box, preventing it from escaping. The potential energy in this model is given as

$$V(x) = \begin{cases} 0, & 0 < x < L, \\ \infty, & \text{otherwise,} \end{cases}$$

where L is the length of the box and x is the position of the particle within the box.

2.6 Wave functions

In quantum mechanics, the wave function gives the most fundamental description of the behavior of a particle; the measurable properties of the particle (such as its position, momentum and energy) may all be derived from the wave function. The wave function $\psi(x,t)$ can be found by solving the Schrödinger equation for the system

$$i\hbar \frac{\partial}{\partial t} \psi(x,t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x,t) + V(x)\psi(x,t),$$

where $\hbar$ is the reduced Planck constant, m is the mass of the particle, i is the imaginary unit and t is time.

Inside the box, no forces act upon the particle, which means that the part of the wave function inside the box oscillates through space and time with the same form as a free particle:

$$\psi(x, t) = [A\sin(kx) + B\cos(kx)]e^{-i\omega t},$$ (1)

where A and B are arbitrary complex numbers. The frequency of the oscillations through space and time are given by the wavenumber k and the angular frequency ω respectively. These are both related to the total energy of the particle by the expression

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

which is known as the dispersion relation for a free particle. Here one must notice that now, since the particle is not entirely free but under the influence of a potential (the potential V described above), the energy of the particle given above is not the same thing as $\frac{p^2}{2m}$ where p is the momentum of the particle, and thus the wavenumber k above actually describes the energy states of the particle, not the momentum states (id est, it turns out that the momentum of the particle is not given by $p = \hbar k$). In this sense, it is quite dangerous to call the number k a wavenumber, since it is not related to momentum like "wavenumber" usually is. The rationale for calling k the wavenumber is that it enumerates the number of crests that the wavefuntion has inside the box, and in this sense it is a wavenumber. This in discrepancy can be seen more clearly below, when we find out that the energy spectrum of the particle is discrete (only discrete values of energy is allowed) but the momentum spectrum is continuous (momentum can vary continuously) and in particular, the relation $E = \frac{p^2}{2m}$ for the energy and momentum of the particle does not hold. As said above, the reason this relation between energy and momentum does not hold is that the particle is not free, but there is a potential V in the system, and the energy of the particle is $E = T + V$, where T is the kinetic and V the potential energy.
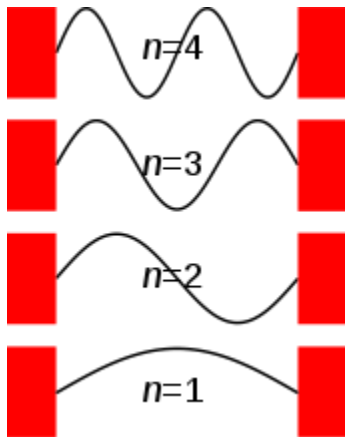
Initial wavefunctions for the first four states in a one-dimensional particle in a box

The size (or amplitude) of the wavefunction at a given position is related to the probability of finding a particle there by $P(x, t) = |\psi(x, t)|^2$. The wave function must therefore vanish

everywhere beyond the edges of the box. Also, the amplitude of the wavefunction may not "jump" abruptly from one point to the next.[1] These two conditions are only satisfied by wave functions with the form

$$\psi_n(x, t) = \begin{cases} A\sin(k_n x)e^{-i\omega_n t}, & 0 < x < L, \\ 0, & \text{otherwise,} \end{cases}$$

where n is a positive integer. Usually in quantum mechanics it is also demanded that the derivative of the wavefuntion in addition to the wavefuntion itself be continuous; here this demand would lead to the only solution being the constant zero function, which is not what we desire, so we give up this demand (as this system with infinite potential can be regarded as a nonphysical abstract limiting case, we can treat it as such and "bend the rules"). Note that giving up this demand means that the wavefuntion is not a differentiable function at the boundary of the box, and thus it can be said that the wavefuntion does not solve the Schrödinger equation at the boundary points $x = 0$ and $x = L$ (but does solve everywhere else). The wavenumber above is restricted to certain, specific values given by

$$k_n = \frac{n\pi}{L}, \quad \text{where} \quad n = \{1, 2, 3, 4, \ldots\},$$

where L is the size of the box. Negative values of n are neglected, since they give wave functions identical to the positive n solutions except for a physically unimportant sign change. Here one sees that only a discrete set of energy values and wavenumbers k are allowed for the particle

2.7 Energy levels
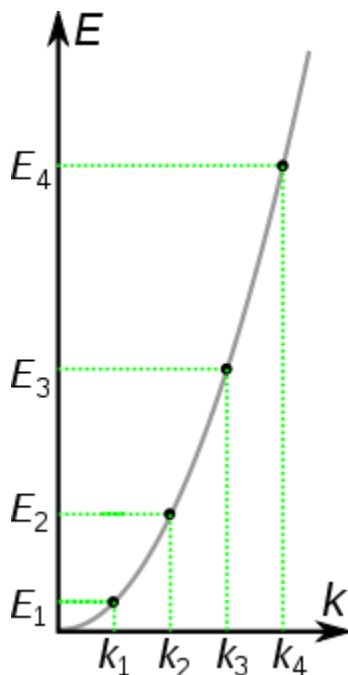
The energy of a particle in a box (black circles) and a free particle (grey line) both depend upon wavenumber in the same way. However, the particle in a box may only have certain, discrete energy levels.

The energies which correspond with each of the permitted wavenumbers may be written as

$$E_n = \frac{n^2 \hbar^2 \pi^2}{2mL^2} = \frac{n^2 h^2}{8mL^2}.$$

The energy levels increase with n^2, meaning that high energy levels are separated from each other by a greater amount than low energy levels are. The lowest possible energy for the particle (its *zero-point energy*) is found in state 1, which is given by[1]

$$E_1 = \frac{\hbar^2 \pi^2}{2mL^2}.$$

The particle, therefore, always has a positive energy. This contrasts with classical systems, where the particle can have zero energy by resting motionlessly. This can be explained in terms of the uncertainty principle, which states that the product of the uncertainties in the position and momentum of a particle is limited by

$$\Delta x \Delta p \geq \frac{\hbar}{2}$$

It can be shown that the uncertainty in the position of the particle is proportional to the width of the box. Thus, the uncertainty in momentum is roughly inversely proportional to the width of the box The kinetic energy of a particle is given by $E = p^2/(2m)$, and hence the minimum kinetic energy of the particle in a box is inversely proportional to the mass and the square of the well width, in qualitative agreement with the calculation above.

2.8 Position and momentum

In classical physics, the particle can be detected anywhere in the box with equal probability. In quantum mechanics, however, the probability density for finding a particle at a given position is derived from the wavefuntion as $P(x) = |\psi(x)|^2$. For the particle in a box, the probability density for finding the particle at a given position depends upon its state, and is given by

$$P_n(x) = \begin{cases} \frac{2}{L} \sin^2 \left(\frac{n\pi x}{L} \right); & 0 < x < L \\ 0; & \text{otherwise.} \end{cases}$$

Thus, for any value of n greater than one, there are regions within the box for which $P(x) = 0$, indicating that *spatial nodes* exist at which the particle cannot be found.

In quantum mechanics, the average, or expectation value of the position of a particle is given by

$$\langle x \rangle = \int_{-\infty}^{\infty} x P_n(x) \, dx.$$

For the steady state particle in a box, it can be shown that the average position is always $\langle x \rangle = L/2$, regardless of the state of the particle. For a superposition of states, the expectation value of the position will change based on the cross term which is proportional to $\cos(\omega t)$.

The variance in the position is a measure of the uncertainty in position of the particle:

$$\text{Var}(x) = \int_{-\infty}^{\infty} (x - \langle x \rangle)^2 P_n(x) \, dx = \frac{L^2}{12} \left(1 - \frac{6}{n^2 \pi^2} \right)$$

The probability density for finding a particle with a given momentum is derived from the wavefuntion as $P(x) = |\phi(x)|^2$. As with position, the probability density for finding the particle at a given momentum depends upon its state, and is given by

$$P_n(p) = \frac{2\pi L}{\hbar} \frac{n^2 \left(1 - (-1)^n \cos(kL) \right)}{\left(k^2 L^2 - \pi^2 n^2 \right)^2}$$

where, again, $k = p/\hbar$. The expectation value for the momentum is then calculated to be zero, and the variance in the momentum is calculated to be:

$$\text{Var}(p) = \left(\frac{\hbar n \pi}{L} \right)^2$$

The uncertainties in position and momentum (Δx and Δp) are defined as being equal to the square root of their respective variances, so that:

$$\Delta x \Delta p = \frac{\hbar}{2} \sqrt{\frac{n^2 \pi^2}{3} - 2}$$

This product increases with increasing n, having a minimum value for $n=1$. The value of this product for $n=1$ is about equal to $0.568\,\hbar$ which obeys the Heisenberg uncertainty principle, which states that the product will be greater than or equal to $\hbar/2$

The idealized situation of a particle in a box with infinitely high walls is an application of the Schrodinger equation which yields some insights into particle confinement. The wave function must be zero at the walls and the solution for the wave function yields just sine waves.

The longest wavelength is

$$\lambda = 2L$$

and the higher modes have wavelengths given by

$$\lambda = \frac{2L}{n} \quad \text{where} \quad n = 1, 2, 3, 4...$$

When this is substituted into the DeBroglie relationship it yields momentum

$$p = \frac{h}{\lambda} = \frac{nh}{2L}$$

When the momentum expression for the particle in a box

$$p = \frac{h}{\lambda} = \frac{nh}{2L} \quad \text{where} \quad n = 1, 2, 3, 4...$$

is used to calculate the energy associated with the particle

$$\frac{1}{2}mv^2 = \frac{p^2}{2m} = \frac{n^2h^2}{8mL^2} = E_n \qquad \text{Energy for nth quantum state for a particle in an infinite box}$$

Though oversimplified, this indicates some important things about bound states for particles:

1. The energies are quantized and can be characterized by a quantum number n
2. The energy cannot be exactly zero.
3. The smaller the confinement, the larger the energy required.

If a particle is confined into a rectangular volume, the same kind of process can be applied to a three-dimensional "particle in a box", and the same kind of energy contribution is made from each dimension. The energies for a three-dimensional box are

$$E = \frac{(n_1^2 + n_2^2 + n_3^2)h^2}{8mL^2}$$

This gives a more physically realistic expression for the available energies for contained particles. This expression is used in determining the density of possible energy states for electrons in solids.

Chapter Three Statistical Thermodynamics

3.1 Introduction

All that we have learnt in thermodynamics are the behavior of *macroscopic systems*, either from a scientific or engineering viewpoint. Examples of such systems might include a piston–cylinder assembly, a heat exchanger, or a battery. Typically, the analysis of macroscopic systems uses conservation or field equations related to classical mechanics, thermodynamics, or electromagnetics. One have learnt to understand that classical thermodynamics is inherently limited in its ability to explain the behavior of even the simplest thermodynamic system. The reason for this deficiency rests with its inadequate treatment of the atomic behavior underlying the gaseous, liquid, or solid states of matter. Without proper consideration of constituent *microscopic systems*, such as a single atom or molecule, it is impossible for the practitioner to understand fully the evaluation of thermodynamic properties, the meaning of thermodynamic equilibrium, or the influence of temperature on transport properties such as the thermal conductivity or viscosity. Developing this elementary viewpoint is the purpose of a course in statistical thermodynamics. As you will see, such fundamental understanding is also the basis for creative applications of classical thermodynamics to macroscopic devices.

3.2 A Classification Scheme for Statistical Thermodynamics
The framework of statistical thermodynamics can be divided into three conceptual themes. The first is equilibrium statistical thermodynamics with a focus on independent particles. Here, we assume no intermolecular interactions among the particles of interest; the resulting simplicity permits excellent a priori calculations of macroscopic behavior. Examples include the ideal gas, the pure crystalline metal, and blackbody radiation. The second theme is again equilibrium statistical thermodynamics, but now with a focus on dependent particles. In this case, intermolecular interactions dominate as, for example, with real gases, liquids, and polymers. Typically, such intermolecular interactions become important only at higher densities; because of

the resulting mathematical difficulties, calculations of macroscopic properties often require semi-empirical procedures.

The third conceptual theme might be labeled nonequilibrium statistical thermodynamics. Here, we are concerned with the dynamic behavior that arises when shifting between different equilibrium states of a macroscopic system. Although time-correlation methods presently constitute an active research program within nonequilibrium statistical thermodynamics, the focus is on those dynamic processes that can be linked to basic kinetic theory. As such, we will explore the molecular behavior underlying macroscopic transport of momentum, energy, and mass. In this way, kinetic theory can provide a deeper understanding of the principles of fluid mechanics, heat transfer, and molecular diffusion. Nonequilibrium statistical thermodynamics also provides an important path for the understanding and modeling of chemical kinetics, specifically, the rates of elementary chemical reactions. Statistical thermodynamics will allow us to calculate from atomic and molecular properties the thermodynamic properties of ideal gases, real gases, and metals. Examples might include equations of state, measurable properties such as specific heats and the internal energy, and also ephemeral properties such as the entropy and free energies. The beauty of statistical thermodynamics is that it reveals our natural world. It will allow one to appreciate the limitations of classical thermodynamics, in particular, the first, second, and third laws of thermodynamics.

Chapter 4 **Probability and Statistics**

In preparation for our study of statistical thermodynamics, we first review some fundamental notions of probability theory, with a special focus on those statistical concepts relevant to atomic and molecular systems.

4.1 Probability: Definitions and Basic Concepts

Probability theory is concerned with predicting statistical outcomes. Simple examples of such outcomes include observing a head or tail when tossing a coin, or obtaining the numbers 1, 2, 3, 4, 5, or 6 when throwing a die. For a fairly-weighted coin, we would, of course, expect to see a head for 1/2 of a large number of tosses; similarly, using a fairly weighted die, we would expect to get a four for 1/6 of all throws. We can then say that the probability of observing a head on one toss of a fairly-weighted coin is 1/2 and that for obtaining a four on one throw of a fairly-weighted die is 1/6. This heuristic notion of probability can be given mathematical formality via the following definition:

Given *Ns mutually exclusive*, equally likely points in *sample space*, with *Ne* of these points corresponding to the *random event A*, then the probability

$$P(A) = Ne/Ns \ .$$

Here, *sample space* designates the available *Ns* occurrences while *random event A* denotes the subset of sample space given by *Ne* ≤ *Ns*. The phrase *mutually exclusive* indicates that no two outcomes can occur simultaneously in a single sample space; this criterion is obviously required if we are to convert our heuristic understanding of chance to a well defined mathematical probability.

As a further example, for a standard deck of playing cards, we have 52 points in sample space, of which four represent aces. Hence, the probability of drawing a single ace from a well-mixed deck is

$$P(A)=4/52=1/13,$$

where the event *A* designates the random drawing of an ace. Visually, the relation between event *A* and sample space can be described by a so-called Venn diagram, as shown in Fig. 4.1. Here, sample points resulting in event *A* fall within the area *A*, while those not resulting in event *A* fall elsewhere in the surrounding box, whose total area represents the entire sample space. Hence, assuming a uniform point density, we find that the ratio of the cross-hatched area to the total area in Fig. 4.1 provides a visual representation of $P(A)$. Similarly, from the viewpoint of set theory, we observe **t**hat for a fairly-weighted die the random event of obtaining an even number $E = \{2, 4, 6\}$ from within the entire sample space $S = \{1, 2, 3, 4, 5, 6\}$ clearly occurs with probability $P(A) = 1/2$.
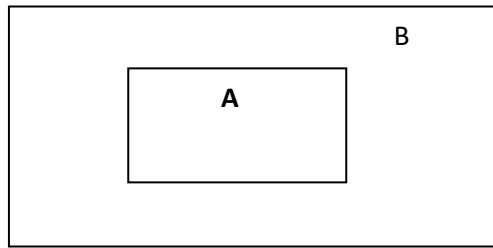
Our notion of probability becomes more complicated if we consider two different random events, *A* and *B*, which can both occur within a given sample space. On this basis, we may define the *compound probability*, $P(AB)$, which represents events *A and B*,

Figure 4.1 Venn diagram representing that portion of sample space which corresponds to random event *A*.

and also the *total probability*, $P(A+B)$, which represents events *A or B* (including both). From the viewpoint of set theory, $P(AB)$ is called the *intersection* of *A* and $B(A \cap B)$, while $P(A+B)$ is labeled the *union* of *A* and *B* ($A \cup B$). Pictorial displays of the (a) intersection and (b) union of *A* and *B* are given by the two Venn diagrams shown in Fig. 4.2. If the events *A* and *B* are mutually exclusive, a single trial by definition permits no overlap in sample space.
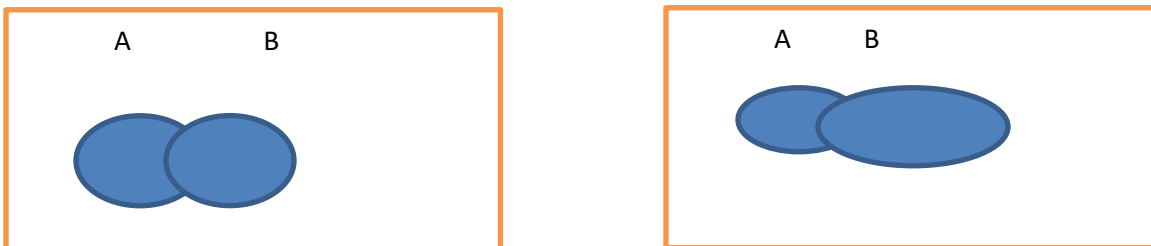

Figure 2.2 Venn diagrams representing (a) $P(AB)$ and (b) $P(A + B)$.

Therefore, $P(AB) = 0$ so that)

$$P(A+B) = P(A) + P(B), \quad (4.1)$$

as displayed by the Venn diagram of Fig. 4.3(a). As an example, the probability of picking a king (*K*) or a queen (*Q*) from a single deck of playing cards is given by the total probability $P(K + Q) = P(K) + P(Q) = 2/13$. In comparison, the probability of picking a king from one deck and a queen from a different deck is $P(KQ) = (1/13)2$. In the latter case, we have two different sample spaces, as indicated by the Venn diagram of Fig. 4.3(b), so that the events are now *mutually independent*. Hence, in general, the compound probability becomes:

$$P(AB) = P(A) \cdot P(B). \quad (4.2)$$

In summary, Eq. (4.1) defines two mutually exclusive events within a single sample space, while Eq. (4.2) defines two mutually independent events within two different sample
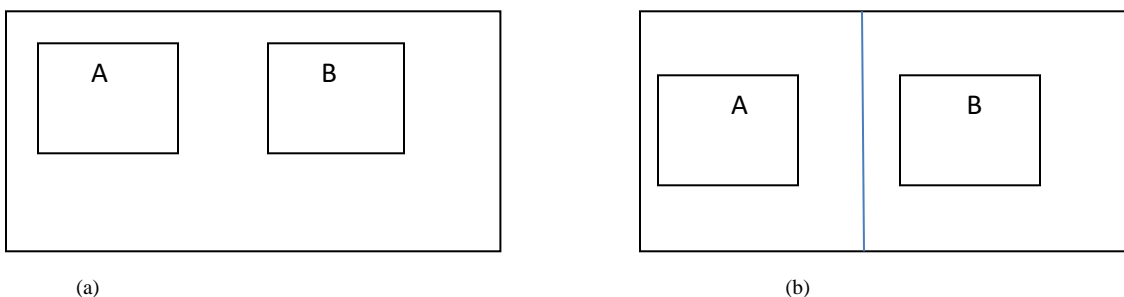

(a)   (b)

Figure 2.3 Venn diagrams describing (a) mutually exclusive and (b) mutually independent events.

spaces. Equations (2.1) and (2.2) can, of course, be extended to more than two events, e.g., $P(A + B + C) = P(A) + P(B) + P(C)$ Mutual Exclusivity (4.3)
$P(ABC) = P(A) \cdot P(B) \cdot P(C)$ Mutual Independence. (4.4)

EXAMPLE 4.1

Five people are arranged in a row at random. What are the probabilities that two particular people will be (a) next to each other and (b) separated by one person between them?

Solution
We first recognize that randomly arranging two previously chosen people with three other people in a row is no different than randomly choosing these same two people after the arrangement. Because choosing two people at random from among five available people represents two mutually independent events, the compound probability with no further information is $(1/5)(1/4)$. However, if we now specify that these two people are either next to each other or one person apart, we must account for the fact that there are many ways of achieving either specification, each of which will enhance the previously unconstrained compound probability. As for many probability analyses, a combination of visual and conceptual approaches often constitutes the most fruitful tactic for solving the problem.

(a) Visualization indicates that for five people in a row, four possible pairs of people can exist next to each other. Conceptually, the persons comprising each pair can also be switched, thus giving eight independent ways of obtaining two people next to each other among five people in a row. Hence, the final probability that two people will be next to each other when five people are arranged in a row at random must be:
$$(1/5)(1/4)(8) = 2/5.$$
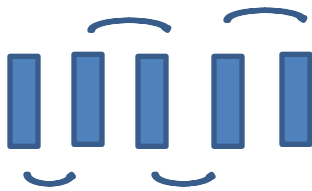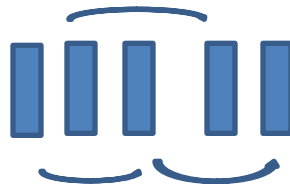
(a)   (b)

Possible pairs of people for (a) two people next to each other and (b) two people one person apart.

(b) Similarly, for two people separated by another person, a combined visual and conceptual analysis gives a final probability of $(1/5)(1/4)(3)(2) = 3/10$. Here, three pairs of people are possible one person apart and the individuals comprising each pair can again be switched. Suppose instead that the five people are arranged in a circle. You should be able to convince yourself that the probability for two people to be either next to each other or separated by another person is now always $1/2$.

4.2 Permutations and Combinations

We now apply probability theory to a sequence of distinguishable objects. Consider, for Example, a basket containing four marbles labeled A, B, C, and D, respectively. Our aim is to randomly select marbles from the basket without replacement. The first marble chosen can be any of four possibilities, the second can be any of the three remaining possibilities, the third chosen must be one of the two remaining possibilities, and the fourth can only be one possibility. Hence, the number of ways that the four sequential but independent choices can be made must be:

$$4 \cdot 3 \cdot 2 \cdot 1 = 24.$$

These 24 possible ways of randomly selecting the four original marbles can be taken as the number of possible arrangements or *permutations* of any single sequence of the four marbles, e.g., $ACDB$. If, on the other hand, the marbles were not labeled, then the 24 possible rearrangements would be irrelevant as the marbles would be indistinguishable. In this case, the 24 permutations would become only one *combination*. Moreover, only a single collection or combination of the four marbles would exist, even if labeled, if we simply chose to disregard any ordering of the random objects.

This distinction between permutations and combinations can be pursued further by considering the number of ways by which we can choose M items from a sample of N available objects without replacement, in one case including and in the other case excluding the effect of labeling or ordering. The objects, for example, could be M marbles chosen from a basket containing N marbles, or M cards chosen from a deck of N cards. Following the procedure outlined in the previous paragraph, the number of permutations is $P(N,M) = N(N-1) \cdots (N-M+1)$ or

$$P(N,M) = N! / (N-M)! , \quad (4.5)$$

which is defined as the number of permutations of N objects taken M at a time. We note, by the way, that $P(N,M) = N!$ When $M = N,$ so that Eq. (4.5) requires that we define $0! = 1$.

In comparison, the number of combinations represents all the different subsets containing M items that can be sampled from N distinct objects. Here, the particular arrangement of M objects within a subset is irrelevant; thus, the number of combinations can be obtained from the number of permutations of Eq. (2.5) via division by the number of permutations, $M!$, for the subset of M distinct objects. Hence,

$$C(N,M) = P(N,M)/M! \text{ or } C(N,M) = N!/ (N - M)! M! \quad (4.6)$$

which is defined as the number of combinations of N objects taken M at a time. We note that $C(N,M)$ can also be interpreted as the number of different arrangements of N objects when M of these objects are of one distinct type and $(N - M)$ are of a second type. This is the interpretation of $C(N,M)$.

EXAMPLE 2.2

You wish to choose three marbles at random from an urn containing four marbles labeled A, B, C, and D.
(a) Determine the number of permutations and combinations for the above scenario.
(b) Identify explicitly each combination and permutation for the three chosen marbles.
Solution
(a) The number of permutations when letting $N = 4$ and $M = 3$ is

$$P(N,M) = N!/ (N - M)! = 4!/1! = 24.$$

Similarly, the number of combinations is:

$$C(N,M) = N!/(N - M)! M! = 4!/1! 3! = 4.$$

(b) The four combinations are ABC, ABD, ACD, and BCD. Each of the four combinations can be permuted $3! = 6$ ways, for a total of 24 permutations. Consider, for example, the ABC combination, which offers the following six permutations: ABC, ACB, BAC, BCA, CAB, and CBA.

4.3 Probability Distributions: Discrete and Continuous

You are no doubt familiar with the concept of a grade distribution as a way of reporting results for a course examination. Consider, for example, the simplified distribution of test scores shown for a class of 20 students in Table 2.1. If we convert the number of students associated with each test score to an appropriate fraction of students, we obtain

Table 2.1 Simplified grade distribution

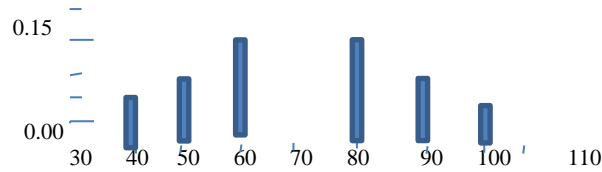| Number of students	Fraction of Students	Test scores
1	0.05	100
2	0.10	90
4	0.20	80
6	0.30	70
4	0.20	60
2	0.10	50
1	0.05	40
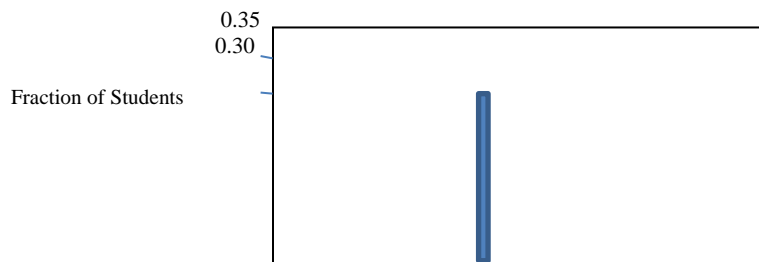
Figure 4.4 Histogram representing Table 2.1.

the *histogram* shown in Fig. 4.4. This histogram is an example of a *discrete probability distribution*, for which

$$\sum_i P(x_i) = 1. \quad (4.7)$$

In other words, the number of students obtaining each grade, x_i, has been *normalized* so that the sum over all probabilities or fractional numbers of students, $P(x_i)$, is unity. For any discrete distribution, the *mean* may be defined as

$$\overline{x} = (x_i) \equiv P(x_i) \, x_i \quad (4.8)$$

while the *variance* or mean-square deviation is

$$\sigma^2 = (x_i - \overline{x})^2, \quad (4.9)$$

where both the overbar and the brackets denote an average or expected value. The square root of the variance, σ, is commonly called the standard deviation; it provides a measure of the width for the probability distribution. Expanding Eq. (4.9) leads to

$$\sigma^2 = (x^2 - 2x_i \, \overline{x_i} + \overline{x^2})$$

so that the variance can be expressed as the mean of the square minus the square of the mean. Hence, the standard deviation can be written as

$$\sigma = \overline{x^2} - \overline{x}^{\,2} \quad (2.10)$$

where

$$\overline{x^2} = (x^2_i) = P_i \, \Sigma(x_i) \, x^2 \quad (4.11)$$

We now consider a more realistic distribution of test scores for a group of 400 students rather than 20 students. In this case, we might expect Table 2.1 to contain all possible integer grades between 40 and 100. Hence, the histogram of Fig. 4.4 would approach a more continuous distribution, as displayed by the *probability density function*, $f(x)$, in Fig. 4.5. Normalization would now be given by:

$$\int f(x)dx = 1$$
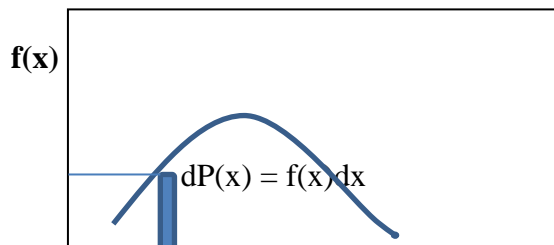
dx   x

Figure 2.5 Continuous distribution functions.

when integrated over all possible values of x. Therefore, the probability density function itself does *not* represent a probability; rather, the probability must be evaluated from knowledge of $f(x)$ via integration. As an example, the probability of achieving values of x between a and b would be obtained from

$$P(a \leq x \leq b) = \int_a^b f(x)\, dx.$$

4.4 The Binomial Distribution

The binomial distribution is of fundamental importance in probability theory, as it describes quite simply any sequence of experiments having two possible outcomes. As an example, consider the tossing of an unfairly-weighted coin, for which the two outcomes are either a head or tail. Suppose that the probability of obtaining a head is p, while that for a tail is $q = 1 - p$. Now, for a sequence of N tosses, the probability of M heads and $(N - M)$ tails in a particular sequence is $p^M q^{N-M}$, as each toss is an independent event in a new sample space. However, because M heads and $(N - M)$ tails can be achieved in more than one way, we must determine the number of possible sequences of heads and tails if we wish to determine the final probability. But the number of possible sequences is just the number of ways N total objects can be arranged into M identical objects of one type and $(N - M)$ identical objects of a second type. This description defines the number of combinations of N objects taken M at a time, $C(N,M)$, as specified by Eq. (4.6). Hence, the probability of tossing M heads and $(N - M)$ tails regardless of order becomes

$$B(M) = C(N,M)p^M q^{N-M}$$

or

$$B(M) = N!/M!\ (N - M)!\ (p^M(1 - p)^{N-M}),\ (4.14)$$

where $B(M)$ represents the well-known binomial probability distribution. This discrete distribution can be interpreted in many different ways. For example, the probabilities p and $(1 - p)$ can indicate the chances of success and failure or right and left steps of a random walk, as well as of heads and tails in the tossing of a coin. Therefore, in general, N always represents the total number of repeated trials in any binary sequence.

EXAMPLE 4.3

Determine the probability that, in six throws of a fairly-weighted die, the side with four pips will land upright at least twice.

Solution

The probability of landing a four on any throw is $1/6$ (success) and thus the probability of not landing a four on any throw is $5/6$ (failure). Consequently, the probability that four pips will not appear ($M = 0$) in six throws ($N = 6$) must be

$B(0) = 6!/0!\ (6 - 0)!(1/6)^0(5/6)^6 = 0.335.$

Similarly, the probability that four pips will appear once ($M = 1$) in six throws is

$B(1) = 6!/1!\ (6 - 1)!(1/6)(5/6)^5 = 0.402.$

As $B(0)$ and $B(1)$ represent mutually exclusive events, the probability from Eq. (4.3) that four pips will appear at least twice in a sequence of six throws must be

$P(M \geq 2) = 1 - P(M < 2) = 1 - [B(0) + B(1)]$

or

$P(M \geq 2) = 1 - [0.335 + 0.402] = 1 - 0.737 = 0.263.$

4.5 The Poisson Distribution

While the binomial distribution holds for any finite number of repeated trials, physical processes involving large numbers of particles, such as in statistical thermodynamics, imply $N \to \infty$. For such circumstances, the binomial distribution can be simplified to two more familiar distributions, one discrete and the other continuous. We now proceed with these simplifications by first assuming $p \to 0$, which we will find leads to the Poisson distribution. This distribution is particularly applicable to photon-counting processes, for which the total number of photons counted, $N \to \infty$, while the possibility of observing any single photon, $p \to 0$.

We begin by expressing the binomial distribution, Eq. (4.14), as :

$$B(M) = N(N-1) \cdots (N-M+1)/M! \, (\mu/N)^M (1-p)^{N-M} \quad (4.14)$$

,

where the mean $\mu \equiv M = Np$ from Eq. (4.15). We then have

$\lim B(M) = N^M / M! \, (\mu/N)^M (1-p)^N = \mu^M/M! ((1-p)^{\mu/p}. \quad (4.15),$

From the fundamental mathematical definition of the quantity, $e = 2.71828$, it can be shown that $\lim(1-p)^{1/p} = e-1$.
$\qquad p \to 0$
Hence, for $N \to \infty$ and $p \to 0$, the binomial distribution becomes the discrete Poisson distribution,

$$P(M) = e^{-\mu} \mu^M/M! \quad (4.15b)$$

Because $P(M)$ is based on $B(M)$, the standard deviation for the Poisson distribution can be obtained from Eq. 4.16) by invoking $p \to 0$, thus giving

$\sigma = Np = \sqrt{\mu}, \qquad (4.18)$

as can also be demonstrated from direct application of Eq. (4.10).We thus find, from Eq. (4.18), that a greater mean value implies a broader range of expected outcomes when the physical system of interest follows Poisson statistics. Employing Eq. (4.15), we also note that $P(M+1)/P(M) = \mu/(M+1)$

,

which indicates a rapid drop in probability for the Poisson distribution as $M \to \infty$. Nevertheless, the Poisson distribution generally remains a good approximation to the binomial distribution for $\mu = Np << \sqrt{N}$.

4.6 Euler–Maclaurin Summation Formula

It often proves convenient in statistical mechanics to approximate the summation of a function over a discrete variable by an analogous integration over a continuous variable. The accuracy of this procedure can be assessed by implementation of the Euler–Maclaurin summation formula. Consider a function $f(n)$ that changes only gradually with increasing integer values, $0 \leq n \leq \infty$. If, in addition,

$\lim f(n) = 0$

$n \to \infty$

,

we may show via contour integration in the complex plane (Hecht, 1990) that

∞

$\Sigma f(n) = \int_0^\infty f(n)\, dn + \frac{1}{2} f(0)\ 1/12 f'(0) + 1/720 f'''0) + \ldots\ldots$

$n=0$

where $f(0)$ is $f(n)$ evaluated at $n = 0$, $f'(0)$ is its first derivative evaluated at $n = 0$, and

$f'''(0)$ is its third derivative evaluated at $n = 0$.

4.7 The Gaussian Distribution.

When the number of trials $N \to \infty$, but p is not small, the binomial distribution becomes the continuous Gaussian distribution rather than the discrete Poisson distribution ($p \to 0$). The Gaussian distribution is particularly applicable to various diffusive processes, for which the total number of molecules $N \to \infty$.

The familiar form of the continuous Gaussian distribution,
$$G(z) = 1/\sqrt{2\pi}\sigma \exp{-z^2/2} \qquad (4.16)$$
.where $z = (x - \mu)/\sigma$

Note that $G(z)$ is symmetrical about z because of its dependence on $z2$, unlike many cases for the discrete binomial or Poisson distributions. Equation (4.16) also indicates that the peak value for $G(z)$ is always $1/\sqrt{2\pi}\sigma$. In general, the Gaussian distribution can be shown to be a satisfactory approximation to the binomial distribution if both $Np \geq 5$ and $Nq \geq 5$.

4.8 Combinatorial Analysis for Statistical Thermodynamics

Quantum mechanics ultimately predicts discrete *energy levels* for molecular systems. As we will see later, each such level is actually composed of a finite number of allowed *energy states*. The number of energy states per energy level is called the *degeneracy*. For our purposes, we can model each energy level of energy, *ej,* as an independent bookshelf holding a specified number of baskets equal to the value of the degeneracy, *gj* , as shown in Figure 4.1 below The height of each individual shelf represents its energy. The equivalent containers denote potential storage locations for the molecules of the thermodynamic system at each energy level. For statistical purposes, we will eventually need to know the distribution of molecules among these energy states,. We now move toward this

goal by considering the number of ways that N objects (molecules) can be placed in M containers (energy states) on a single shelf (energy level). Before we can make such combinatorial calculations, however, we must introduce two other important features of quantum mechanics
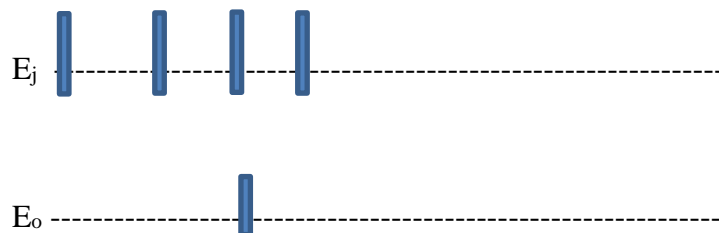

Figure 4.1 Bookshelf model for energy level E_j with degeneracy $gj = 4$.

Chapter 5 Essential Concepts from Quantum Mechanics

In preparation for our study of elementary statistical thermodynamics, we recapitulate and expand somewhat on those essential notions from quantum mechanics. To avoid unnecessary complications, the background needed to derive or fully understand the following four conceptual presumptions is discussed. The first concept is that energy for a single atom or molecule is always *quantized*, as implied by the three discrete energy levels designated $\varepsilon 0$, $\varepsilon 1$, and $\varepsilon 2$ in the energy-level diagram of Fig. 5.1. As discussed previously, quantization is suggested by the discrete lines appearing in both atomic and molecular spectra. The second essential concept is that the available energy levels are not necessarily equally probable. This statement is in accord with the different number of energy states associated with each energy level, as described by the degeneracies, $g0$, $g1$, and $g2$, in Fig. 5.1. We will find out that each energy or *quantum* state is defined by its own unique set of so-called *quantum numbers*.
This unique specification at either the atomic or molecular level suggests that each energy state, rather than each energy level, can be considered equally likely

E_2 ☐ ☐ ☐ ☐ ☐ $g_2 = 5$

E_1 ☐ $g_1 = 1$
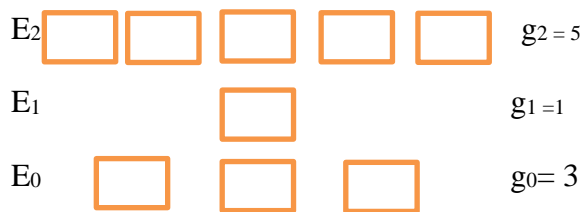
E_0 ☐ ☐ ☐ $g_0 = 3$

Figure 5.1 Example of Energy diagram

The third essential concept is that most particles are indistinguishable rather than distinguishable owing to the probabilistic nature of matter at atomic dimensions. From a different perspective, the uncertainty principle suggests that molecules can be counted but not discerned because their momentum and position cannot be specified simultaneously. If, however, atoms are aligned structurally either within or on the surface of a crystalline material, for example, then discernment of particular atoms becomes possible, thus making the particles distinguishable. The fourth and final concept concerns the possible number of particles permitted per energy state. For some particles, such as protons and electrons, only one particle is allowed per energy state. Other particles, such as photons, display no limit on the number allowed per energy state. For simplicity, the former particles are called *fermions* while the latter are called *bosons*. In summary, then, the crucial information needed from quantum mechanics to make statistical thermodynamic calculations is the energy, εj, and the degeneracy, gj, corresponding to the jth energy level of the relevant atom or molecule. In addition, for statistical purposes, we must know whether the particle of interest is (1) distinguishable or indistinguishable and (2) fermion or boson

5.2 The Ensemble Method of Statistical Thermodynamics
The most general statistical procedure for calculating thermodynamic properties is called the ensemble method, as developed by the American engineer and scientist J. Willard Gibbs (1839–1903). The Gibbs approach works for both dependent and independent particles,

thus making it very powerful for realistic thermodynamic systems. The more restrictive Maxwell–Boltzmann (M–B) method, which presumes an *isolated system of independent particles is more suitable to model* as if they were composed of independent particles. Examples include not only the ideal gas, but also electrons, radiation, and the crystalline solid. In addition, pedagogically, the M–B method is more intuitive and thus it furnishes a necessary foundation for fully appreciating the rigorous beauty of the Gibbs method.

"An ***ensemble*** is a theoretical collection of a very large number η of systems, each of which replicates the *macroscopic* thermodynamic system under investigation."

There are three main types of ensembles, depending on what type of macroscopic system is being replicated to create the ensemble. The ***microcanonical*** ensemble is composed of η *isolated* systems (N, V, U), for which the total number of particles, N, the volume, V, and the internal energy, U, are the replicated thermodynamic properties. The ***canonical*** ensemble is composed of η *closed, isothermal* systems (N, V, T), for which the total number of particles, the volume, and the temperature, T, are the independent thermodynamic parameters. Finally, the ***grand canonical*** ensemble is composed of η *open, isothermal* systems (μ, V, T), for which the chemical potential, μ, the volume, and the temperature are the constant independent variables. In general, an ensemble is really a super system composed of η replicated thermodynamic systems, such as a room, a piston–cylinder assembly, or a nozzle. To ensure proper replication, no mass or energy can be permitted to cross the overall boundary of the constructed super system. Hence, the entire ensemble, no matter what type, must always be isolated

Thermal insulation provides a heat bath of constant temperature to ensure isothermal conditions for each member of the ensemble

Each replicated macroscopic member of the ensemble is a closed, isothermal system at constant N, V, and
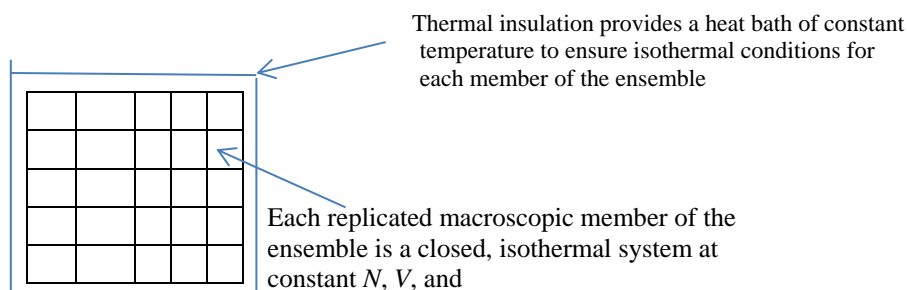
Figure 5.2 Diagram for the canonical ensemble; the entire ensemble is an isolated super system

Fig 5.2 demonstrates the construction of an *isolated supersystem* for the canonical ensemble. We may now link the M–B and Gibbs approaches by first recalling that the M–B method requires an isolated system containing N independent particles. If we consider the canonical ensemble of Fig. 5.2, this condition can be assured by artificially restricting each member of the ensemble to a single particle. In this way, the supersystem of the ensemble method becomes the thermodynamic system of the M–B method. The particles are guaranteed to be independent because each particle is associated with an independent member of the ensemble. The M–B system is guaranteed to be isolated because the constructed supersystem is isolated by definition. The result is that the number of independent particles, N, becomes equal to the number of replicated macroscopic systems, η.

Chapter 6 Statistical Mechanics

Statistical mechanics is devoted to the analysis of the internal motions of a many-particle system using probability theory—a subject area that is known as statistical mechanics.

6.1 Specifications of State of Many-Particle System

Let us first consider how we might specify the state of a many-particle system. Consider the simplest possible dynamical system, which consists of a single spinless particle moving classically in one dimension. Assuming that we know the particle's equation of motion, the state of the system is fully specified once we simultaneously measure the particle's displacement, q, and momentum, p. In principle, if we know q and p then we can calculate the state of the system at all subsequent times using the equation of motion. In practice, it is impossible to specify q and p exactly, because there is always an intrinsic error in any experimental measurement. Consider the time evolution of q and p. This can be visualized by plotting the point (q, p) in the q-p plane. This plane is generally known as phase-space. In general, as time progresses, the point (q, p) will trace out some very complicated pattern in phase-space. Suppose that we divide phase-space into rectangular cells of uniform dimensions δq and δp. Here, δq is the intrinsic error in the position measurement, and δp the intrinsic error in the momentum measurement. The area of each cell is

$\delta q \delta p = h_0$, (6.1) where h_0 is a small constant having the dimensions of angular momentum.

The coordinates q and p can now be conveniently specified by indicating the cell in phase-space into which they plot at any given time. This procedure automatically ensures that we do not attempt to specify q and p to an accuracy greater than our experimental error, which would clearly be pointless. Let us now consider a single spinless particle moving in three dimensions. In order to specify the state of the system, we now need to know three q-p pairs: that is, q_x-p_x, q_y-p_y, and q_z-p_z. Incidentally, the number of q-p pairs needed to specify the state of the system is usually called the number of degrees of freedom of the system. Thus, a single particle moving in one dimension constitutes a one degree of freedom system, whereas a single particle moving in three dimensions constitutes a three degree of freedom system. Consider the time evolution of q and p, where q = (q_x, q_y, q_z), et cetera. This can be visualized by plotting the point (q, p) in the six-dimensional q-p phase-space. Suppose that we divide the q_x-p_x plane into rectangular cells of uniform dimensions δq and δp, and do likewise for the q_y-p_y and q_z-p_z planes. Here, δq and δp are again the intrinsic errors in our measurements of position and momentum, respectively. This is equivalent to dividing phase-space up into regular six-dimensional cells of volume h_0^3. The coordinates q and p can now be conveniently specified by indicating the cell in phase-space into which they plot at any given time.

In principle, we can specify the state of the system to arbitrary accuracy, by taking the limit $h_0 \rightarrow 0$. In reality, we know from Heisenberg's uncertainty principle that it is impossible to simultaneously measure a coordinate, q_i, and its conjugate momentum, p_i, to greater accuracy than $\delta q_i \delta p_i = h/2$. Here, h is Planck's constant divided by 2π. This implies that

$$h_0 \geq h/2. \quad (6.2)$$

In other words, the uncertainty principlesets a lower limit on how finely we can chop up classical phase-space. In quantum mechanics, we can specify the state of the system by giving its wavefunction at timet,

$$\psi(q1,\cdots,qf, s1,\cdots, sg,t), \quad (6.3)$$

where f is the number of translational degrees of freedom and g the number of internal (e.g.spin) degrees of freedom. For instance, if the system consists of N spin-one-half particles then there will be 3N translational degrees of freedom, and N spin degrees of freedom (becausethespinofeachparticlecaneitherbedirecteduapordownalongthez-axis). Alternatively, if the system is in a stationary state (i.e.an eigen state of the Hamiltonian) then we can just specify f g quantum numbers. Either way, the future time evolution of the wave function is fully determined by Schrodinger's equation. In reality, this approach is not practical because the Hamiltonian of the system is only known approximately. Typically, we are dealing with a system consisting of many weakly-interacting particles. We usually know the Hamiltonian for completely non-interacting particles, but the component of the Hamiltonian associated with particle interactions is either impossibly complicated, or not very well known. We can define approximate stationary eigenstates using the Hamiltonian for non-interacting particles. The state of the system is then specified by the quantum numbers identifying these eigenstates. In the absence of particle interactions, if the system starts off in a stationary state then it stays in that state forever, so its quantum numbers never change. The interactions allow the system to make transitions between different "stationary" states, causing its quantum numbers to change in time.

6.3 H-Theorem
Consider a system of weakly-interacting particles. In quantum mechanics, we can write the Hamiltonian for such a system as $H = H_0 + H_1$, (6.4)
where H_0 is the Hamiltonian for completely non-interacting particles, and H_1 a small correction due to the particle interactions.

We can define approximate stationary eigenstates of the system using H_0.

Thus, $H_0 \Psi r = Er \Psi r$, (6.5) where the index r labels a state of energy Er and eigen state Ψr. In general, there are many different eigenstates with the same energy—these are called degenerate states. (See Section C.10.) For example, consider N non-interacting spin less particles of mass m confined in a cubic box of dimension L. According to standard wave-mechanics, the energy levels of the ith particle are given by
$$ei = h^2\pi^2/ 2mL^2 (n_{1i}^2 + n_{2i}^2 + n_{3i}^2), \quad (6.6)$$
where n_{1i}, n_{2i}, andn$_{3i}$ are three (positive integer) quantum numbers.

The overall energy of the system is the sum of the energies of the individual particles, so that for a general state r,
$$Er = \sum_{i=1,N} e_i, \quad (6.7)$$

The overall state of the system is thus specified by 3N quantum numbers(i.e.three quantum numbers per particle). There are clearly very many different arrangements of these quantum numbers that give the same overall energy.

Consider, now, a statistical ensemble of systems made up of weakly-interacting particles. Suppose that this ensemble is initially very far from equilibrium. For instance, the systems in the ensemble might only be distributed over a very small subset of their accessible states. If each system starts off in a particular stationary state (i.e., with a particular set of quantum numbers) then, in the absence of particle interactions, it will remain in that state forever. Hence, the ensemble will always stay far from equilibrium, and the principle of equal apriori probabilities will never be applicable. In reality, particle interactions cause each system in the ensemble to make transitions between its accessible "stationary" states. This allows the overall state of the ensemble to change in time.

Small interactions between particles cause transitions between the approximate stationary states of the system. Thus, there exists some transition probability per unit time, W_{rs}, that a system originally in state r ends up in state, s, as a result of these interactions. Likewise, there exists a probability per unit time, W_{sr}, that a system in state, s, makes a transition to state r. These transition probabilities are meaningful in quantum mechanics provided that the particle interaction strength is sufficiently small, there is an early continuous distribution of accessible energy levels, and we consider time intervals that are not too small. These conditions are easily satisfied for the types of systems usually analyzed via statistical mechanics (e.g., nearly-ideal gases). One important conclusion of quantum mechanics is that the forward and backward transition probabilities between two states are the same, so that $W_{rs} = W_{sr}$

6.4 Relaxation Time
The H-theorem guarantees that an isolated many-particle system will eventually reach an equilibrium state, irrespective of its initial state. The typical time required for this process to take place is called the relaxation time, and depends, in detail, on the nature of the inter-particle interactions. The principle of equal a priori probabilities is only valid for equilibrium states. It follows that we can only safely apply this principle to systems that have remained undisturbed for many relaxation times since they were setup, or last interacted with the outside world. The relaxation time for the air in a typical classroom is very much less than one second. This suggests that such air is probably in equilibrium most of the time, and should, therefore, be governed by the principle of equal a priori probabilities. In fact, this is known to be the case. Consider another example. Our galaxy, the "Milky Way," is an isolated dynamical system made up of about 1011 stars. In fact, it can be thought of as a self-gravitating "gas" of stars. At first sight, the Milky Way would seem to be an ideal system on which to test out the ideas of statistical mechanics. Stars in the Milky Way interact via occasional near-miss events in which they exchange energy and momentum. Actual collisions are very rare indeed. Unfortunately, such interactions take place very infrequently, because there is a lot of empty space between the stars. The best estimate for the relaxation time of the Milky Way is about 10^{13} years. This should be compared with the estimated age of the Milky Way, which is only about 10^{10} years. It is clear that, despite its great age, the Milky Way has not been around long enough to reach an equilibrium state. This suggests that the principle of equal a priori probabilities cannot be used to describe stellar

dynamics. Not surprisingly, the observed velocity distribution of the stars in the vicinity of the Sun is not governed by this principle.

6.5 Reversibility and Irreversibility

Previously, we mentioned that, on a microscopic level, the laws of physics are in variant under time reversal. In other words, microscopic phenomena look physically plausible when run in reverse. We usually say that these phenomena are reversible. Does this imply that macroscopic phenomena are also reversible? Consider an isolated many-particle system that starts off far from equilibrium. According to the H-theorem, it will evolve towards equilibrium and as it does so, the macroscopic quantity H will decrease. But, if we run this process backwards in time then the system will appear to evolve away from equilibrium, and the quantity H will increase. This type of behavior is not physical because it violates the H-theorem. In other words, if we saw a film of a macroscopic process then we could very easily tell if it was being run backwards. For instance, suppose that by some miracle, we were able to move all of the oxygen molecules in the air in some classroom to one side of the room, and all of the nitrogen molecules to the opposite side. We would not expect this state to persist for very long. Pretty soon the oxygen and nitrogen molecules would start to intermingle, and this process would continue until they were thoroughly mixed together throughout the room. This, of course, is the equilibrium state for air. In reverse, this process appears completely unphysical. We would start off from perfectly normal air, and suddenly, for no good reason, the air's constituent oxygen and nitrogen molecules would appear to separate, and move to opposite sides of the room. This scenario is not impossible, but, from everything we know about the world around us, it is spectacularly unlikely. We conclude, therefore, that macroscopic phenomena are generally irreversible, because they appear unphysical when run in reverse. How does the irreversibility of macroscopic phenomena arise? It certainly does not come from the fundamental laws of physics, because these laws are all reversible. In the previous example, the oxygen and nitrogen molecules intermingled by continually scattering off one another. Each individual scattering event would look perfectly reasonable viewed in reverse. However, the net result of this scattering event appears unphysical when run backwards. How can we obtain an irreversible process from the combined effects of very many reversible processes? This is a vitally important question. Unfortunately, we are not quite at the stage where we can formulate a convincing answer. Note, however, that the essential irreversibility of macroscopic phenomena is one of the key results of statistical thermodynamics.

Chapter 7 Heat and Work

7.1 Brief History of Heat and Work

Antoine Lavoisier

In 1789, the French scientist Antoine Lavoisier published a famous treatise on chemistry which, among other things, demolished the then prevalent theory of combustion. This theory, known to history as the phlogiston theory, is so extraordinary stupid that it is not even worth describing. In place of phlogiston theory, Lavoisier proposed the first reasonably sensible scientific interpretation of heat. Lavoisier pictured heat as an invisible, tasteless, odorless, weightless fluid,

which he called calorific fluid. He postulated that hot bodies contain more of this fluid than cold bodies. Furthermore, he suggested that the constituent particles of calorific fluid repel one another, causing heat to flow spontaneously from hot to cold bodies when they are placed in thermal contact. The modern interpretation of heat is, or course, somewhat different to Lavoisier's calorific theory. Nevertheless, there is an important subset of problems, involving heat flow, for which Lavoisier's approach is rather useful. These problems often crop up as examination questions. For example: "A clean dry copper calorimeter contains 100 grams of water at 30◦ degrees centigrade. A 10 gram block of copper heated to 60◦ centigrade is added. What is the final temperature of the mixture?". How do we approach this type of problem? According to Lavoisier's theory, there is an analogy between heat flow and incompressible fluid flow under gravity. The same volume of liquid added to containers of different (uniform)cross-sectional area fills them to different heights. If the volume is V, and the cross-sectional area is A, then the height is h = V/A. In a similar manner, the same quantity of heat added to different bodies causes them to rise to different temperatures. If Q is the heat and θ is the (absolute) temperature then θ = Q/C, where the constant C is termed the heat capacity. [This is a somewhat oversimplified example. In general, the heat capacity is a function of temperature, so that C = C(θ).] If two containers, filled to different heights, with a free-flowing incompressible fluid are connected together at the bottom, via a small pipe, then fluid will flow under gravity, from one to the other, until the two heights are the same. The final height is easily calculated by equating the total fluid volume in the initial and final states. Thus,

$h_1 A_1 + h_2 A_2 = hA_1 + hA_2$, (7.1)
giving $h = h_1 A_1 + h_2 A_2 A_1 + A_2$ (7.2)
Here, h_1 and h_2 are the initial heights in the two containers, A_1 and A_2 are the corresponding crosssectional areas, and his the final height. Likewise, if two bodies, initially at different temperatures, are brought into thermal contact then heat will flow, from one to the other, until the two temperatures are the same. The final temperature is calculated by equating the total heat in the initial and final states. Thus, $\theta_1 C_1 + \theta_2 C_2 = \theta C_1 + \theta C_2$, (4.3) giving

$\theta = \theta_1 C_1 + \theta_2 C_2 C_1 + C_2$
, (7.4)

and fluid flow works because, in Lavoisier's theory, heat is a conserved quantity, just like the volume of an incompressible fluid. In fact, Lavoisier postulated that heat was an element. Notethatatomswerethoughttobeindestructiblebeforenuclearreactions were discovered, so the total amount of each element in the cosmos was assumed to be a constant. Thus, if Lavoisier had cared to formulate a law of thermodynamics from his calorific theory then he would have said that the total amount of heat in the universe as a constant. In 1798, Benjamin Thompson, an Englishman who spent his early years in pre-revolutionary America, was minister for war and police in the German state of Bavaria. One of his jobs was to oversee the boring of cannons in the state arsenal. Thompson was struck by the enormous, and seemingly inexhaustible, amount of heat generated in this process. He simply could not understand where all this heat was coming from. According to Lavoisier's calorific theory, the heat must flow into the cannon from its immediate surroundings, which should, therefore, become colder. The flow should also eventually cease when all of the available heat has been extracted. In fact, Thompson observed that the surroundings of the cannon got hotter, not colder, and that the heating process continued unabated as long as the boring machine was operating. Thompson postulated that some of the mechanical work done on the cannon by the boring machine was being converted into heat. At

the time, this was quite a revolutionary concept, and most people were not ready to accept it. This is somewhat surprising, because, by the end of the eighteenth century, the conversion of heat into work, by steam engines, was quite commonplace. Nevertheless, the conversion of work intoheatdidnotgainbroadacceptanceuntil1849, when an English physicist called James Prescott Joule published the results of a long and painstaking series of experiments. Joule confirmed that work could indeed be converted into heat. Moreover, he found that the same amount of work always generates the same quantity of heat. This is true regardless of the nature of the work (e.g., mechanical, electrical, et cetera). Joule was able to formulate what became known as the work equivalent of heat. Namely, that 1 newton meter of work is equivalent to 0.241 calories of heat. A calorie is the amount of heat required to raise the temperature of 1 gram of water by 1 degree centigrade. Nowadays, we measure both heat and work using the same units, so that one newton meter, or joule, of work is equivalent to one joule of heat. In 1850, the German   physicist Clausius correctly postulated that the essential conserved quantity is neither heat nor work, but some combination of the two which quickly became known as energy, from the Greek Energeia meaning "activity" or "action. " According to Clausius, the change in the internal energy ofa macroscopic body can be written

$\Delta E = Q - W$, (4.5)

where Q is the heat absorbed from the surroundings, and W is the work done on the surroundings. This relation is known as the first law of thermodynamics.

7,2 Microscopic Interpretation of Heat and Work

Consider a macroscopic system, A, that is known to be in a given macrostate. To be more exact, consider an ensemble of similar macroscopic systems, A, where each system in the ensemble is in one of the many microstates consistent with the given macrostate. There are two fundamentally different ways in which the average energy of A can change due to interaction with its surroundings. If the external parameters of the system remain constant then the interaction is termed a purely thermal interaction. Any change in the average energy of the system is attributed to an exchange of heat with its environment. Thus, $\Delta E = Q$, (4.7) where Q is theheat absorbed by the system. On a microscopic level,the energies oftheindividual microstates are unaffected by the absorption of heat. In fact, it is the distributionof the systems in the ensembleoverthevarious microstates that is modified. SupposethatthesystemAisthermallyinsulatedfromitsenvironment. Thiscanbeachievedby surrounding it by anadiabaticenvelope(i.e.,anenvelopefabricatedoutofamaterialthatisapoor conductor of heat, such a fiber glass). Incidentally, the term adiabatic is derived from the Greek adiabatos, which means "impassable." In scientific terminology, an adiabatic process is one in whichthereisnoexchangeofheat. ThesystemAisstillcapableofinteractingwithitsenvironment via its external parameters. This type of interaction is termed mechanical interaction, and any change in the average energy of the system is attributed to work done on it by its surroundings. Thus, $\Delta E = -W$, (4.8)

46 4.4. Quasi-Static Processes

whereW istheworkdonebythesystemonitsenvironment. Onamicroscopiclevel,theenergyof thesystemchanges becausetheenergies oftheindividualmicrostatesare functionsoftheexternal parameters. [See Equation (4.6).] Thus, if the external parameters are changed then, in general, the energies of all of the systems in the ensemble are modified (because each is in a specific

microstate). Such a modification usually gives rise to a redistributionof the systems in the ensemble overtheaccessiblemicrostates(withoutany heat exchangewiththeenvironment). Clearly, from a microscopicviewpoint,performing work on amacroscopic systemis quitea complicatedprocess. Nevertheless,macroscopicworkisaquantitythatiseasy tomeasureexperimentally. Forinstance, ifthesystem A exertsaforceFonitsimmediatesurroundings,andthechangein externalparameters corresponds to a displacement x of the center of mass of the system, then the work done by A on its surroundingsis simply $W = F \cdot x$: (4.9) that is, the product of the force and the displacement along the line of action of the force. In a general interaction of the system A with its environment there is both heat exchange and work performed. We can write $Q \equiv \Delta E + W$, (4.10) which serves as the general definition of the absorbed heat Q. (Hence, the equivalence sign.) The quantityQissimplythechangeinthemeanenergyofthesystemthatisnotduetothemodification of the external parameters. Note that the notion of a quantity of heat has no independent meaning apart from Equation (4.10). The mean energy, E, and work performed, W, are both physical quantitiesthat can be determined experimentally,whereas Q is merely a derived quantity.