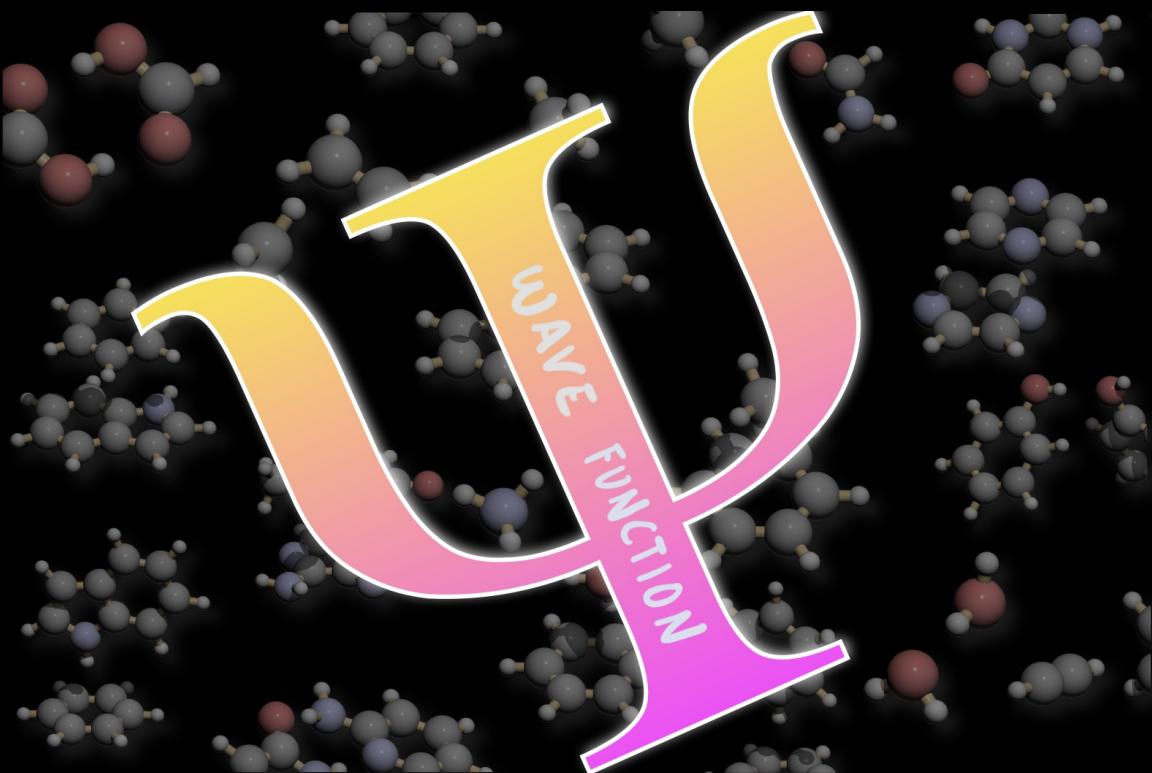


The Live Textbook of Physical Chemistry

Volume 2: Quantum Mechanics



Roberto Peverati

Contents

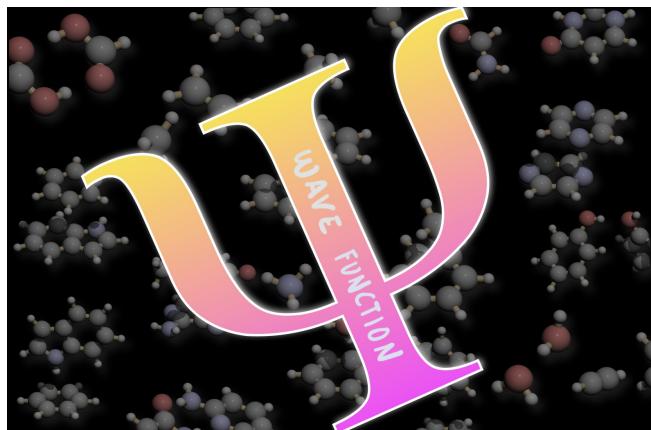
Preface	vi
How to use this book	vii
1 Introduction to Quantum Mechanics	1
1.1 Historical Context	1
1.2 The Ultraviolet Catastrophe	2
1.3 The Photoelectric Effect	4
1.4 Wave-Particle Duality	6
1.5 Chapter Review	7
2 Brief Summary of Classical Physics	11
2.1 Classical Mechanics	11
2.2 Coulomb's Law	18
2.3 Chapter Review	20
3 The Schrödinger Equation	25
3.1 The Time-Independent Schrödinger Equation	25
3.2 The Time-Dependent Schrödinger Equation	27
3.3 Solution for the Free Particle	28
3.4 Chapter Review	30

4 Analytically Solvable Models	35
4.1 The Particle in a Box	35
4.2 The Particle on a Ring	40
4.3 The Particle on a Sphere	47
4.4 The Harmonic Oscillator	57
4.5 Chapter Review	60
5 The Hydrogen Atom	65
5.1 Wave Function Analysis	67
5.2 Chapter Review	80
6 Operators and Mathematical Background	85
6.1 Operators in Quantum Mechanics	85
6.2 Eigenfunctions and Eigenvalues	89
6.3 Common Operators in Quantum Mechanics	91
6.4 Chapter Review	94
7 Spin	99
7.1 Stern-Gerlach Experiment	99
7.2 Sequential Stern-Gerlach Experiments	101
7.3 Spin Operators	102
7.4 Chapter Review	104
8 Postulates of Quantum Mechanics	109
8.1 Postulate 1: The Wave Function Postulate	110
8.2 Postulate 2: Experimental Observables	110
8.3 Postulate 3: Individual Measurements	110
8.4 Postulate 4: Expectation Values and Collapse of the Wave-function	111

8.5 Postulate 5: Time Evolution	112
8.6 Postulate 6: Pauli Exclusion Principle	112
8.7 Chapter Review	114
9 Quantum Weirdness	119
9.1 The double-slit experiment	119
9.2 Heisenberg's Uncertainty Principle	122
9.3 Tunneling	124
9.4 Chapter Review	127
10 Many-Electron Atoms	133
10.1 Many-Electron Wave Functions	133
10.2 Approximated Hamiltonians	136
10.3 Chapter Review	140
11 Introduction to Molecules	145
11.1 The Molecular Hamiltonian	145
11.2 The Born-Oppenheimer Approximation	146
11.3 Solving the Electronic Eigenvalue Problem	148
11.4 Chapter Review	150
12 The Chemical Bond in Diatomic Molecules	155
12.1 The Chemical Bond in the Hydrogen Molecular Cation	155
12.2 The Chemical Bond in the Hydrogen Molecule	158
12.3 Chapter Review	162
13 The Chemical Bond in Polyatomic Molecules	167
13.1 The Chemical Bond in the Water Molecule Using a Minimal Basis	167
13.2 Hartree-Fock Calculation for Water	169

13.3 Shapes and Energies of Molecular Orbitals	175
13.4 Chapter Review	178
14 Spectroscopy	183
14.1 Rotational Spectroscopy	184
14.2 Vibrational Spectroscopy	187
14.3 Electronic Spectroscopy	190
14.4 Chapter Review	192

Preface



This textbook is the official textbook for the Physical Chemistry 2 Course (CHM 3002) at Florida Tech.

The instructor for this course and author of this textbook is Dr. Roberto Peverati.

Contacts: rpeverati@fit.edu, Office: OPS 333, (321) 674-7735

Chemistry Program, Department of Biomedical and Chemical Engineering and Science Florida Institute of Technology, Melbourne, FL.

This live open textbook is distributed under the CC-BY-SA 4.0 License and it was funded by the Florida Tech Open Educational Resources Grant Program: A Collaboration of the Teaching Council, eEducation, and the Evans Library.

How to use this book

If you have taken P-Chem 1 at Florida Tech in the Fall Semester 2020, you should already be familiar with this format. Everything we did with the Live Textbook of P-Chem 1 in FS2020 will be happening this semester with P-Chem 2. Please read this book carefully, since everything that will be in your exams is explained here. Since this book is specifically tailored for the CHM 3002 course at Florida Tech, there are no superfluous parts. In other words, everything in it might be subject to question in the quizzes and the final exam.

Definitions and exercises are usually numbered and are highlighted in the text in this format (lighter grey, indented, and following a grey vertical bar). Please study the definitions carefully since they are fundamental concepts that will be used several times in the remainder of the text, and they will be subject to quizzes and exams. Exercises are essential for cementing the concepts, and you should attempt to execute them first without looking at the solution. Even if you were able to solve an exercise on your own, always read the solution after, since it might contain additional explanations expanding the main concepts in the text.

Navigating the book should be straightforward. On each page, there is a useful sidebar on the left that gives you an overview of all chapters, and a toolbar at the top with important tools. Arrows to shift between chapters might also be present, depending on your browser. If you are old-school and prefer a pdf, you can download a printout by clicking on the toolbar's corresponding icon. If you are *really* old-school and prefer a printed book, the best solution is to download the pdf and print it yourself. It is a LaTeX book, and I can promise you it will look good on paper. However, I cannot provide physical copies to each student. In the toolbar, you will find a useful search box that is capable of searching the entire book. The most adventurous will find in the toolbar a link to the raw GitHub source code. Feel free to head on over there and fork the book.

Each chapter of this book represents one week of work in the classroom and at home. The sidebar on the left will reflect your syllabus, as well as the main structure of the class on Canvas. The book is a live document, which means it will be updated throughout the semester with new material. While you are not required to check it every day, you might want to review each week's chapter before the lecture on Friday.

If you spot a mistake or a typo, contact Dr. Peverati via email and you will receive a credit of up to three points towards your final score, once the typo has been verified and corrected.

Chapter 1

Introduction to Quantum Mechanics

1.1 Historical Context

Quantum mechanics is a branch of physics that deals with the behavior of matter and energy at the atomic and subatomic level. It is a fundamental theory that has revolutionized our understanding of the physical world, and it is a cornerstone of modern chemistry. But the road to the development of quantum mechanics was a long and winding one that spanned several decades and involved many of the greatest minds in physics.

The story of quantum mechanics begins in the late 19th century, with the discovery of the electron by J.J. Thomson in 1897. This discovery, along with the discovery of radioactivity by Marie and Pierre Curie in the same year, sparked a flurry of activity in the field of atomic physics. Scientists were eager to understand the structure and behavior of atoms, and they soon discovered that classical physics was inadequate for this task.

In 1900, the German physicist Max Planck made a breakthrough that would lay the groundwork for quantum mechanics. Planck was studying the spectrum of light emitted by a heated object, and he found that the energy of the emitted radiation was not continuous, as classical physics would have predicted, but instead came in discrete packets or “quanta”. This discovery was revolutionary and opened up a new era in physics.

The next major milestone in the development of quantum mechanics came

in 1913, with the Bohr model of the atom. Danish physicist Niels Bohr proposed a new model of the atom that incorporated the idea of quantized energy levels. According to Bohr's model, electrons could only occupy certain discrete energy levels, and when they moved from one level to another, they emitted or absorbed energy in the form of light. Bohr's model was a significant step forward, but it had limitations. It could only accurately predict the behavior of simple atoms with one or two electrons, and it did not account for the complex spectra of more complex atoms.

It was not until the mid-1920s that a truly comprehensive theory of quantum mechanics began to emerge. In 1925, Austrian physicist Erwin Schrödinger published his famous equation, which described the wave-like behavior of particles at the atomic level. Schrödinger's equation was a mathematical triumph, but it was also highly abstract and difficult to interpret. At the same time, Werner Heisenberg, a German physicist, was developing a different approach to quantum mechanics, known as matrix mechanics. Heisenberg's theory was based on matrices, which are arrays of numbers that can be used to describe the properties of quantum particles. Heisenberg's approach was highly mathematical, but it provided a more intuitive interpretation of the behavior of particles at the atomic level. The two approaches, matrix mechanics and wave mechanics, were eventually shown to be mathematically equivalent, and together they formed the basis of modern quantum mechanics.

In the decades that followed, quantum mechanics continued to evolve, with new discoveries and theories expanding our understanding of the behavior of matter and energy at the atomic and subatomic level. Today, quantum mechanics is a thriving field of research, with applications in chemistry, physics, and many other areas of science.

The development of quantum mechanics was driven by a series of discoveries and experiments that challenged classical physics and its inability to explain the behavior of atoms. These are the “ultraviolet catastrophe”, the photoelectric effect, and the wave-particle duality. We will discuss each of these cases below.

1.2 The Ultraviolet Catastrophe

The ultraviolet (UV) catastrophe, also called the Rayleigh–Jeans catastrophe, is the prediction of classical electromagnetism that the intensity of the radiation emitted by an ideal black body at thermal equilibrium goes to

infinity as wavelength decreases (see figure 1.1¹).

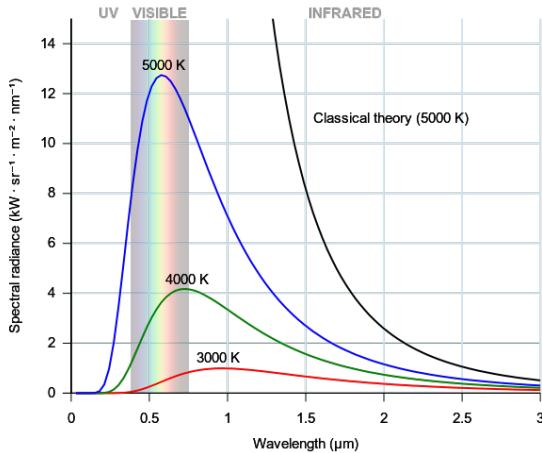


Figure 1.1: The ultraviolet catastrophe is the error at short wavelengths in the Rayleigh–Jeans law for the energy emitted by an ideal black body. The error, much more pronounced for short wavelengths, is the difference between the Rayleigh–Jeans law —black—and Planck’s law—blue.

A black body is an idealized object that absorbs and emits all frequencies. Classical physics can be used to derive an approximated equation describing the intensity of a black body radiation as a function of frequency for a fixed temperature. The result is known as the Rayleigh–Jeans law, which for wavelength λ , is:

$$I(T) = \frac{2ck_B T}{\lambda^4} \quad (1.1)$$

where I is the intensity of the radiation —expressed as the power emitted per unit emitting area, per steradian, per unit wavelength (spectral radiance)— c is the speed of light, k_B is the Boltzmann constant, and T is the temperature in kelvins. The paradox —or rather the breakdown of the Rayleigh–Jeans formula— happens at small wavelength λ . If we take the limit for $\lambda \rightarrow 0$ in eq. (1.1), we obtain that $I \rightarrow \infty$. In other words, as the wavelength of the emitted light gets smaller (approaching the UV range),

¹This picture is taken from Wikipedia by user Darth Kule, and is in the Public Domain

the intensity of the radiation approaches infinity, and the black body emits an infinite amount of energy. This divergence for low wavelength (high frequencies) is called the ultraviolet catastrophe, and it is clearly unphysical.

Max Planck explained the black body radiation in 1900 by assuming that the energies of the oscillations of the electrons responsible for the radiation must be proportional to integral multiples of the frequency:

$$E = nh\nu = nh\frac{c}{\lambda}, \quad (1.2)$$

with n being an integer number bigger than zero. Planck's assumptions led to the correct form of the spectral function for a black body:

$$I(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/(\lambda k_B T)} - 1}. \quad (1.3)$$

If we now take the limit for $\lambda \rightarrow 0$ of eq. (1.3), it is easy to prove that I goes to zero, in agreement with the experimental results, and our intuition. Planck also found that for $h = 6.626 \times 10^{-34}$ J s, the experimental data could be reproduced exactly. Nevertheless, Planck could not offer a good justification for his assumption of energy quantization. Physicists did not take this energy quantization idea seriously until Einstein invoked a similar assumption to explain the photoelectric effect.

1.3 The Photoelectric Effect

The photoelectric effect refers to the phenomenon where electrons are emitted from a metal surface when it is illuminated with light of a certain frequency. This effect was first observed by Heinrich Hertz in 1887 and later explained by Albert Einstein in 1905, who proposed that light energy was quantized into discrete particles, or photons, each with an energy proportional to its frequency.

Before Einstein's explanation, the classical prediction for the photoelectric effect was based on the wave theory of light, which assumed that light was a continuous wave. According to this assumption, the intensity of light is associated with the energy of the electromagnetic waves that make up the light. The intensity of light, I , is defined, as in the previous case, using the spectral radiance, which is the power emitted per unit emitting area, per steradian, per unit wavelength:

$$I = \frac{P}{A}, \quad (1.4)$$

where P is the power of the light, and A is the area of the wave front. The power of the light is related to the energy of the electromagnetic waves that make up the light, and is given by:

$$P = \frac{dE}{dt}, \quad (1.5)$$

where P is the power of the light, E is the energy of the electromagnetic waves, and t is time. Therefore, the intensity of light is related to the energy of the electromagnetic waves that make up the light, and can be increased by increasing the energy of the waves or by increasing the number of waves per unit area (i.e., increasing the wave's amplitude or its frequency).

Applying this classical wave theory of light to explain the photoelectric effect, we deduce that when light shines on a metal surface, the energy of the light is transferred to the electrons in the metal continuously. The electrons would then be emitted when they accumulate enough energy to escape the surface. Since the energy of the light increases with its intensity, the kinetic energy of the emitted electron, K , should be proportional to the intensity of the incident light, I_{in} :

$$K = c_{\text{PE}} I_{\text{in}}, \quad (1.6)$$

with c_{PE} being a constant that depends on the metal and the characteristics of the incident light. In other words, if we change the intensity of the incident light, the kinetic energy of the emitted electrons would change accordingly. This explanation predicts also that a sufficiently dim light should result in a delayed emission. This prediction, however, was contradicted by Hertz's experimental observations, which showed no dependence on the light's intensity and no delayed emission (results were clearly independent on the duration of the exposure). In all of his experiments, electrons were dislodged from the surface of the metal only when the light exceeded a certain frequency, regardless of the intensity of the light.

Einstein used the lack of dependency on the intensity of the light in Hertz's experiment to propose that a beam of light is not a wave propagating through space, but a swarm of discrete energy packets, known as photons.

Each of this photons will have an energy that is given by the smallest possible frequency in Planck's theory of the black body (corresponding to $n = 1$ in eq. (1.2)):

$$E = h\nu. \quad (1.7)$$

The energy of the incident photons is therefore quantized, and only photons with a frequency greater than the threshold frequency (corresponding to the work function of the metal) can eject electrons from the metal surface. A low-frequency beam at a high intensity will not build up the energy required to produce electrons, because there is no process for accumulation. The maximum kinetic energy that the electron can gain, K_{\max} , is equal to the energy of the incident photon, $E_{\text{in}} = h\nu_{\text{in}}$, minus the work function of the metal ϕ , which is the minimum amount of energy required to remove an electron from the metal surface. Mathematically, this can be expressed as:

$$K_{\max} = E_{\text{in}} - \phi = h\nu_{\text{in}} - \phi. \quad (1.8)$$

1.4 Wave-Particle Duality

Einstein had shown that the momentum of a photon is

$$p = \frac{h}{\lambda}. \quad (1.9)$$

This can be easily shown as follows. Assuming $E = h\nu$ for a photon and $\lambda\nu = c$ for an electromagnetic wave, we obtain

$$E = \frac{hc}{\lambda} \quad (1.10)$$

Now we use Einstein's relativity result, $E = mc^2$, and the definition of momentum $p = mc$, to find:

$$\lambda = \frac{h}{p}, \quad (1.11)$$

which is equivalent to eq. (1.9). Note that m refers to the relativistic mass, not the rest mass, since the rest mass of a photon is zero. Since light can behave both as a wave (it can be diffracted, and it has a wavelength), and

as a particle (it contains packets of energy $h\nu$), de Broglie reasoned in 1924 that matter also can exhibit this wave-particle duality. He further reasoned that matter would obey the same eq. (1.11) as light. In 1927, Davisson and Germer observed diffraction patterns by bombarding metals with electrons, confirming de Broglie's proposition.

Rewriting the previous equations in terms of the wave vector, $k = \frac{2\pi}{\lambda}$, and the angular frequency, $\omega = 2\pi\nu$, we obtain the following two equations

$$\begin{aligned} p &= \hbar k \\ E &= \hbar\omega, \end{aligned} \tag{1.12}$$

which are known as **de Broglie's equations**. We will use those equation to develop wave mechanics in the next chapters.²

1.5 Chapter Review

1.5.1 Study Questions

1. The “ultraviolet catastrophe” refers to which failure of classical physics?

- a. Inability to explain discrete atomic emission spectra
 - b. Prediction that black-body intensity vanishes at long wavelengths
 - c. Prediction that black-body intensity diverges at short wavelengths
 - d. Failure to explain the existence of photons
 - e. Prediction that all objects emit only in the ultraviolet region
-

2. What key assumption did Planck introduce to resolve the ultraviolet catastrophe?

²This chapter has been written with help from ChatGPT

- a. Charge is quantized in units of the electron charge
 - b. Energy of oscillators is quantized in multiples of $h\nu$
 - c. Momentum is continuous but position is quantized
 - d. Light speed depends on wavelength in vacuum
 - e. Temperature of a black body is quantized
-

3. What is the main qualitative difference between the Rayleigh–Jeans law and Planck's law at short wavelengths?

- a. Rayleigh–Jeans diverges, Planck's law goes to zero
 - b. Both diverge to infinity but at different rates
 - c. Rayleigh–Jeans is constant, Planck's law oscillates
 - d. Rayleigh–Jeans is valid only in the ultraviolet, Planck only in infrared
 - e. Planck's law diverges while Rayleigh–Jeans goes to zero
-

4. In the classical wave picture of the photoelectric effect, how is the kinetic energy K of emitted electrons predicted to depend on the incident light?

- a. Proportional to the square of the wavelength
 - b. Independent of intensity, depends only on work function
 - c. Proportional to the inverse of exposure time
 - d. Proportional to the intensity of the light
 - e. Independent of both intensity and frequency
-

5. Which experimental observation about the photoelectric effect contradicts the classical wave prediction?

- a. Increase of emission with higher temperature of the metal
 - b. Emission of electrons only for polarized light
 - c. Existence of negative ions at the surface
 - d. Disappearance of current at very high intensities
 - e. Existence of a threshold frequency independent of intensity
-

6. Using $E = h\nu$ and $\lambda\nu = c$, which expression correctly relates photon energy to wavelength?

- a. $E = hc/\lambda$
 - b. $E = h\lambda/c$
 - c. $E = c/(h\lambda)$
 - d. $E = hc\lambda$
 - e. $E = \lambda^2/hc$
-

7. Which de Broglie relation connects particle momentum p with wave vector k ?

- a. $p = k/\hbar$
 - b. $p = \hbar/k$
 - c. $p = \hbar k^2$
 - d. $p = k^2/\hbar$
 - e. $p = \hbar k$
-

8. What is the main conceptual implication of wave–particle duality for matter?

- a. Matter is purely a continuous wave
 - b. Matter has only discrete allowed velocities
 - c. Matter exhibits both wave-like and particle-like properties depending on the experiment
 - d. Matter exists only as photons at small scales
 - e. Matter cannot be localized in space
-

9. In the historical development described in this chapter, which of the following came first chronologically?

- a. Planck's black-body quantization
 - b. Schrödinger's wave equation
 - c. De Broglie's matter waves
 - d. Heisenberg's matrix mechanics
 - e. Bohr's model of the atom
-

10. Which of the following best summarizes the crisis that led from classical to quantum physics as outlined in this Chapter?

- a. Classical theories failed to conserve energy in chemical reactions
 - b. Classical theories could not explain black-body spectra, photoelectric thresholds, and matter diffraction
 - c. Classical mechanics could not describe planetary motion
 - d. Classical thermodynamics predicted negative absolute temperatures
 - e. Classical optics could not describe refraction at dielectric interfaces
-

Answers: 1.c, 2.b, 3.a, 4.d, 5.e, 6.a, 7.e, 8.c, 9.a, 10.b

Chapter 2

Brief Summary of Classical Physics

Quantum mechanics cannot be derived from classical mechanics, but classical mechanics can inspire quantum mechanics. Quantum mechanics is richer and more sophisticated than classical mechanics. Quantum mechanics was developed during the period when physicists had rich knowledge of classical mechanics. In order to better understand how quantum mechanics was developed in this environment, it is better to understand some fundamental concepts in classical mechanics. Classical mechanics can be considered as a special case of quantum mechanics. We will review some classical mechanics concepts here.

2.1 Classical Mechanics

2.1.1 Newtonian Formulation

Classical mechanics as formulated by Isaac Newton (1652-1727) is all about forces. Newtonian mechanics works well for problems where we know the forces and have a reasonable coordinate system. In these cases, the net force acting on a system at position \mathbf{r} in Cartesian 3-dimensional space $\{x, y, z\}$ is simply:

$$F_{\text{net}}(\mathbf{r}) = m\ddot{\mathbf{r}} = m \frac{d^2\mathbf{r}}{dt^2}. \quad (2.1)$$

Or, in other words, if we know the net force acting on a system of mass m at position \mathbf{r} at some time t_0 , we can use eq. (2.1) to calculate the position of the system at any future (or past) time. We have completely determined the dynamical evolution of the system.¹

Example 2.1. A ball of mass m is at ground level and tossed straight up from an initial position x_0 with an initial velocity \dot{x}_0 and subject to gravity alone. Calculate the equation of motion for the ball (i.e. where is the ball going to be after some time t ?).²

Solution: Since the only force acting on the ball is gravity, we can use the equation for the gravitational force to start our derivation:

$$F_{\text{gravity}} = -mG, \quad (2.2)$$

with G the usual gravitational constant ($G = 9.8 \text{ m/s}^2$). We can then replace this expression into eq. (2.1), to obtain:

$$\begin{aligned} -mG &= m\ddot{\mathbf{r}} \\ -G &= \ddot{\mathbf{r}} \\ -G &= \frac{d\dot{\mathbf{r}}}{dt}. \end{aligned} \quad (2.3)$$

Working in a 1-dimensional case, we can replace \mathbf{r} with position x to obtain:

$$\begin{aligned} -mG &= m\ddot{x} \\ -G &= \ddot{x} \\ -G &= \frac{d\dot{x}}{dt}. \end{aligned} \quad (2.4)$$

¹Notice that $\mathbf{r} \in \mathbb{R}^3$ is the position vector and $\dot{\mathbf{r}} \in \mathbb{R}^3$ is the velocity vector. As such, all the equation of classical mechanics are vector equation in 3-dimensional space, and not just simple numerical equation. To simplify the picture, we can restrict ourselves to a 1-dimensional space, replace \mathbf{r} with x , and forget the complications of vector algebra, as we do in the examples.

²This example is based on Rhett Allain's blog post that can be found (here)[<https://rhettallain.com/2018/10/31/classical-mechanics-newtonian-lagrangian-and-hamiltonian/>]

which can then be integrated with respect to time, to obtain:

$$\begin{aligned} -G \int_{t=0}^t dt &= \int_{x_0}^{\dot{x}} d\dot{x} \\ \dot{x} &= \dot{x}_0 - Gt \\ \frac{dx}{dt} &= \dot{x}_0 - Gt, \end{aligned} \tag{2.5}$$

which can be further integrated with respect to time, to give:

$$\begin{aligned} \int_{x_0}^x dx &= \int_{t=0}^t \dot{x}_0 dt - G \int_{t=0}^t t dt \\ x &= x_0 + \dot{x}_0 t - \frac{1}{2} G t^2. \end{aligned} \tag{2.6}$$

This final equation is the equation of motion for the ball, from which we can calculate the position of the ball at any time t . Notice how the equation of motion does not depend on the mass of the ball!

Numerical problem: How much time will a ball ejected from a height of 1 m at an initial velocity of 10 m/s take to hit the floor?

We can use the equation of motion obtained above to solve this problem, and obtain for this specific case $t \simeq 2.12$ s.³

The formula of Newtonian mechanics are not the only one we can use to solve a problem in classical mechanics. We have at least two other equivalent approaches to the same problem that might end up being more useful in certain situations.

2.1.2 Lagrangian Formulation

Another way to derive the equations of motion for classical mechanics is via the use of the Lagrangian and the principle of least action. The Lagrangian formulation is obtained by starting from the definition of the Lagrangian of the system:

³Can you write a python program to do this calculation?

$$L = K - V, \quad (2.7)$$

where K is the kinetic energy, and V is the potential energy. Both are expressed in terms of the coordinates $(\mathbf{r}, \dot{\mathbf{r}})$. Notice that for a fixed time, t , \mathbf{r} and $\dot{\mathbf{r}}$ are independent variables, since $\dot{\mathbf{r}}$ cannot be derived from \mathbf{r} alone.

The time integral of the Lagrangian is called the **action**, and is defined as:

$$S = \int_{t_1}^{t_2} L dt, \quad (2.8)$$

which is a functional: it takes in the Lagrangian function for all times between t_1 and t_2 and returns a scalar value. The equations of motion can be derived from the principle of least action,⁴ which states that the true evolution of a system $q(t)$ described by the coordinate q between two specified states $\mathbf{r}_1 = \mathbf{r}(t_1)$ and $\mathbf{r}_2 = \mathbf{r}(t_2)$ at two specified times t_1 and t_2 is a minimum of the action functional. For a minimum point:

$$\delta S = \frac{dS}{d\mathbf{r}} = 0 \quad (2.9)$$

Requiring that the true trajectory $\mathbf{r}(t)$ minimizes the action functional S , we obtain the equation of motion (figure 2.1⁵). This can be achieved applying classical variational calculus to the variation of the action integral S under perturbations of the path $q(t)$, eq. (2.9). The resulting equation of motion (or set of equations in the case of many dimensions) is sometimes also called the Euler—Lagrange equation:⁶

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\mathbf{r}}} \right) = \frac{\partial L}{\partial \mathbf{r}}. \quad (2.10)$$

⁴Sometimes also called principle of stationary action, or variational principle, or Hamilton's principle.

⁵This diagram is taken from Wikipedia by user Maschen, and distributed under CC0 license

⁶The mathematical derivation of the Euler—Lagrange equaiton is rather long and unimportant at this stage. For the curious, it can be found here.

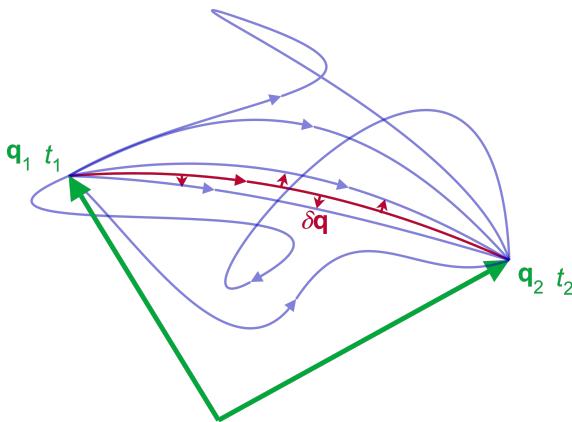


Figure 2.1: Principle of least action: As the system evolves, \mathbf{q} traces a path through configuration space (only some are shown). The path taken by the system (red) has a stationary action under small changes in the configuration of the system.

Example 2.2. Let's apply the Lagrangian mechanics formulas to the same problem as in the previous Example.

The expression of the kinetic energy, the potential energy, and the Lagrangian for our system are:

$$\begin{aligned} K &= \frac{1}{2}m\dot{x}^2 \\ V &= mGx \\ L &= K - V = \frac{1}{2}m\dot{x}^2 - mGx. \end{aligned} \tag{2.11}$$

To get the equation of motion using eq. (2.10), we need to first take the partial derivative of L with respect to x (right hand side):

$$\frac{\partial L}{\partial x} = -mG, \tag{2.12}$$

and then we need the derivative with respect to t of the derivative of the Lagrangian with respect to \dot{x} at the left hand side:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = \frac{d \left(\frac{1}{2} m \dot{x}^2 - m G x \right)}{dt} = m \ddot{x}. \quad (2.13)$$

Putting this together, we get:

$$\begin{aligned} m \ddot{x} &= -mG \\ \ddot{x} &= -G \end{aligned} \quad (2.14)$$

Which is the same result as obtained from the Newtonian method. Integrating twice, we get the exact same formulas that we can use the same way.

The advantage of Lagrangian mechanics is that it is not constrained to use a coordinate system. For example, if we have a bead moving along a wire, we can define the coordinate system as the distance along the wire, making the formulas much simpler than in Newtonian mechanics. Also, since the Lagrangian depends on kinetic and potential energy it does a much better job with constraint forces.

2.1.3 Hamiltonian mechanics

A third way of obtaining the equation of motion is Hamiltonian mechanics, which uses the generalized momentum in place of velocity as a coordinate. The generalized momentum, \mathbf{p} , is defined in terms of the Lagrangian and the coordinates $(\mathbf{r}, \dot{\mathbf{r}})$:

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{r}}}. \quad (2.15)$$

The Hamiltonian is defined from the Lagrangian by applying a Legendre transformation as:⁷

$$H(\mathbf{p}, \mathbf{r}) = \mathbf{p} \dot{\mathbf{r}} - L(\mathbf{r}, \dot{\mathbf{r}}), \quad (2.16)$$

The Lagrangian equation of motion becomes a pair of equations known as the Hamiltonian system of equations:

⁷We have already encountered Legendre transform in The Live Textbook of Physical Chemistry 1 when transforming from the thermodynamic energy to any of the other thermodynamic potentials.

$$\begin{aligned}\dot{\mathbf{p}} &= \frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{r}} \\ \dot{\mathbf{r}} &= \frac{d\mathbf{r}}{dt} = +\frac{\partial H}{\partial \mathbf{p}},\end{aligned}\tag{2.17}$$

where $H = H(\mathbf{p}, \mathbf{r}, t)$ is the Hamiltonian of the system, which often corresponds to its total energy. For a closed system, it is the sum of the kinetic and potential energy in the system:

$$H = K + V.\tag{2.18}$$

Notice the difference between the Hamiltonian, eq. (2.18), and the Lagrangian, eq. (2.7). In Newtonian mechanics, the time evolution is obtained by computing the total force being exerted on each particle of the system, and from Newton's second law, the time evolution of both position and velocity are computed. In contrast, in Hamiltonian mechanics, the time evolution is obtained by computing the Hamiltonian of the system in the generalized momenta and inserting it into Hamilton's equations. This approach is equivalent to the one used in Lagrangian mechanics, since the Hamiltonian is the Legendre transform of the Lagrangian. The main motivation to use Hamiltonian mechanics instead of Lagrangian mechanics comes from the more simple description of complex dynamic systems.

Example 2.3. Let's apply the Hamiltonian mechanics formulas to the same problem in the previous examples.

Using eq. (2.16), the Hamiltonian can be written as:

$$H = m\dot{\mathbf{r}}\dot{\mathbf{r}} - \frac{1}{2}m\dot{\mathbf{r}}^2 + mG\mathbf{r} = \frac{1}{2}m\dot{\mathbf{r}}^2 + mG\mathbf{r}.\tag{2.19}$$

Since the Hamiltonian really depends on position and momentum, we need to get this in terms of \mathbf{r} and \mathbf{p} , with $\mathbf{r} = x$ and $\mathbf{p} = p = m\dot{x}$ for our 1-dimensional system.⁸ Hence we have:

$$H = \frac{p^2}{2m} + mGx,\tag{2.20}$$

⁸These equations are not universally valid, since they depend on the choice of coordinate system.

from which we can use eqs. (2.17) to get:

$$\begin{aligned}\dot{x} &= \frac{\partial H}{\partial p} = \frac{p}{m} \\ \dot{p} &= -\frac{\partial H}{\partial x} = -mG.\end{aligned}\tag{2.21}$$

These equations represent a major difference of the Hamiltonian method, since we describe the system using two first-order differential equations, rather than one second-order differential equation. In order to get the equation of motion, we need to take the derivative of \dot{x} :

$$\ddot{x} = \frac{d}{dt} \left(\frac{p}{m} \right) = \frac{\dot{p}}{m},\tag{2.22}$$

and then replacing the definition of \dot{p} obtained above, we get:

$$\ddot{x} = \frac{-mG}{m} = -G\tag{2.23}$$

which—once again—is the same result obtained for the two previous cases. Integrating this twice, we get the familiar equation of motion for our problem.

2.2 Coulomb's Law

Coulomb's Law states that the magnitude of the electrostatic force of attraction or repulsion between two point charges is directly proportional to the product of the magnitudes of charges and inversely proportional to the square of the distance between them:

$$|\mathbf{F}| = k_e \frac{|q_1 q_2|}{r^2} = \frac{1}{4\pi\epsilon_0} \frac{|q_1 q_2|}{r^2}.\tag{2.24}$$

The constant k_e is called the Coulomb constant and is equal to $1/4\pi\epsilon_0$, where ϵ_0 is the the vacuum permittivity constant; $k_e \simeq 8.988 \times 10^9 \text{ Nm}^2\text{C}^{-2}$. If the product $q_1 q_2$ is positive, the force between the two charges is repulsive; if the product is negative, the force between them is attractive.

2.2.1 Vector form of the Coulomb's Law law

Coulomb's law in vector form states that the electrostatic force \mathbf{F}_1 experienced by a charge, q_1 at position \mathbf{r}_1 , in the vicinity of another charge, q_2 at position \mathbf{r}_2 , in a vacuum is equal

$$\mathbf{F}_1 = \frac{q_1 q_2}{4\pi\epsilon_0} \frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} = \frac{q_1 q_2}{4\pi\epsilon_0} \frac{\hat{\mathbf{r}}_{12}}{|\mathbf{r}_{12}|^2}, \quad (2.25)$$

where $\mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2$ is the vectorial distance between the charges, $\hat{\mathbf{r}}_{12} = \frac{\mathbf{r}_{12}}{|\mathbf{r}_{12}|}$ a unit vector pointing from q_2 to q_1 , and ϵ_0 the vacuum permittivity constant. Here, $\hat{\mathbf{r}}_{12}$ is used for the vector notation.

The vector form of Coulomb's law is simply the scalar definition of the law with the direction given by the unit vector, $\hat{\mathbf{r}}_{12}$, parallel with the line from charge q_2 to charge q_1 . If both charges have the same sign (like charges) then the product $q_1 q_2$ is positive and the direction of the force on q_1 is given by $\hat{\mathbf{r}}_{12}$; the charges repel each other. If the charges have opposite signs then the product $q_1 q_2$ is negative and the direction of the force on q_1 is $-\hat{\mathbf{r}}_{12}$; the charges attract each other.

The electrostatic force \mathbf{F}_2 experienced by q_2 , according to Newton's third law, is $\mathbf{F}_2 = -\mathbf{F}_1$.

2.2.2 Electric field

An electric field is a vector field that associates to each point in space the Coulomb force experienced by a unit test charge. The strength and direction of the Coulomb force \mathbf{F} on a charge q_t depends on the electric field \mathbf{E} established by other charges that it finds itself in, such that $\mathbf{F} = q_t \mathbf{E}$. In the simplest case, the field is considered to be generated solely by a single source point charge. If the field is generated by a positive source point charge q , the direction of the electric field points along lines directed radially outwards from it, i.e. in the direction that a positive point test charge q_t would move if placed in the field. For a negative point source charge, the direction is radially inwards.

The magnitude of the electric field, $|\mathbf{E}|$, can be derived from Coulomb's law, eq. (2.24), by choosing one of the point charges to be the source ($q_1 = q$), and the other to be the test charge ($q_2 = q_t$), and is given by:

$$|\mathbf{E}| = |\mathbf{F}| \cdot \frac{1}{q_t} = \frac{1}{4\pi\epsilon_0} \frac{|q|}{r^2} \quad (2.26)$$

2.2.3 Electric Potential Energy

The electric potential energy of a system of point charges is defined as the work required to assemble this system of charges by bringing them close together, as in the system from an infinite distance.

The electrostatic potential energy, E , of one point charge q at position \mathbf{r} in the presence of an electric field \mathbf{E} is defined as the negative of the work W done by the electrostatic force to bring it from a reference position at infinite distance \mathbf{r}_∞ to that position \mathbf{r} :

$$E = -W_{\mathbf{r}_\infty \rightarrow \mathbf{r}} = - \int_{\mathbf{r}_\infty}^{\mathbf{r}} q \mathbf{E}(\mathbf{r}') \cdot d\mathbf{r}', \quad (2.27)$$

where \mathbf{E} is the electrostatic field and $d\mathbf{r}'$ is the displacement vector in a curve from the reference position \mathbf{r}_∞ to the final position \mathbf{r} .

In the simplest case of two point charges q_1 and q_2 , the electrostatic potential energy at position r , can be derived from Coulomb's Law, eq. (2.24), as:

$$E = |\mathbf{F}| \cdot r = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r}. \quad (2.28)$$

This last equation will become very important in future chapters. Notice the difference between eq. (2.28) for the *energy*, E , and eq. (2.26) for the *magnitude of the electric field*, $|\mathbf{E}|$.

2.3 Chapter Review

2.3.1 Study Questions

1. What is the central quantity in Newtonian mechanics?
 - a. Force
 - b. Action
 - c. Wavefunction

- d. Electric field
 - e. Partition function
-

2. In the ball-under-gravity example, which force is used to set up the equation of motion?

- a. $F_{\text{gravity}} = +mG$
 - b. $F_{\text{gravity}} = -mG$
 - c. $F_{\text{gravity}} = -kx$
 - d. $F_{\text{gravity}} = qE$
 - e. $F_{\text{gravity}} = 0$
-

3. Which statement is true about the equation of motion for the 1D motion of a ball tossed vertically?

- a. It is valid only for very small masses
 - b. It is valid only in two dimensions
 - c. It assumes variable gravitational acceleration
 - d. It requires relativistic corrections
 - e. It does not depend on the mass of the ball
-

4. How is the Lagrangian L of a system defined?

- a. $L = K + V$
 - b. $L = p\dot{r} - H$
 - c. $L = K - V$
 - d. $L = F \cdot r$
 - e. $L = \partial K / \partial t$
-

5. In the Lagrangian formulation, what is the action S ?

- a. The time integral of L between t_1 and t_2
 - b. The instantaneous kinetic energy
 - c. The work done by conservative forces
 - d. The momentum times position
 - e. The time derivative of the Hamiltonian
-

6. In Hamiltonian mechanics, how is the generalized momentum p defined?

- a. $p = m\ddot{r}$
 - b. $p = \frac{\partial H}{\partial r}$
 - c. $p = \frac{\partial r}{\partial V}$
 - d. $p = \frac{\partial L}{\partial \dot{L}}$
 - e. $p = \frac{\partial K}{\partial r}$
-

7. The Hamiltonian $H(p, r)$ is obtained from the Lagrangian via which transformation?

- a. Fourier transformation
 - b. Legendre transformation
 - c. Laplace transformation
 - d. Gauge transformation
 - e. Lorentz transformation
-

8. What is an advantage of Lagrangian mechanics over Newtonian mechanics?

- a. It eliminates the need for forces entirely
 - b. It is valid only for relativistic systems
 - c. It does not require energy conservation
 - d. It works only in one dimension
 - e. It naturally handles constraints and non-Cartesian coordinates
-

9. The magnitude of the electric field $|E|$ created by a point charge q at distance r is given by which of the following formula?

- a. $|E| = \frac{1}{4\pi\epsilon_0} \frac{|q|}{r^2}$
 - b. $|E| = \frac{1}{4\pi\epsilon_0} |q|r^2$
 - c. $|E| = \frac{|q|}{r}$
 - d. $|E| = k_B T/r^2$
 - e. $|E| = \frac{1}{4\pi\epsilon_0} \frac{|q|}{r}$
-

10. The electrostatic potential energy E of two point charges q_1 and q_2 separated by distance r is given by which of the following formula?

- a. $E = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r^2}$
- b. $E = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r}$
- c. $E = q_1 q_2 r$
- d. $E = k_B T r$
- e. $E = \frac{1}{4\pi\epsilon_0} q_1 q_2 r^2$

Answers: 1.a, 2.b, 3.e, 4.c, 5.a, 6.d, 7.b, 8.e, 9.a, 10.b

Chapter 3

The Schrödinger Equation

In 1925, Erwin Schrödinger and Werner Heisenberg independently developed the new quantum theory. Schrödinger's method involves partial differential equations, whereas Heisenberg's method employs matrices; however, a year later the two methods were shown to be mathematically equivalent. Most textbooks begin with Schrödinger's equation, since it seems to have a better physical interpretation via the classical wave equation. Indeed, the Schrödinger equation can be viewed as a form of the wave equation applied to matter waves.

3.1 The Time-Independent Schrödinger Equation

We can start the derivation of the single-particle time-independent Schrödinger equation (TISEq) from the equation that describes the motion of a wave in classical mechanics:

$$\psi(x, t) = \exp[i(kx - \omega t)], \quad (3.1)$$

where x is the position, t is time, $k = \frac{2\pi}{\lambda}$ is the wave vector, and $\omega = 2\pi\nu$ is the angular frequency of the wave. If we are not concerned with the time evolution, we can consider uniquely the derivatives of eq. (3.1) with respect to the location, which are:

$$\begin{aligned}\frac{\partial\psi}{\partial x} &= ik \exp[i(kx - \omega t)] = ik\psi, \\ \frac{\partial^2\psi}{\partial x^2} &= i^2 k^2 \exp[i(kx - \omega t)] = -k^2\psi,\end{aligned}\tag{3.2}$$

where we have used the fact that $i^2 = -1$.

Assuming that particles behaves as wave—as proven by de Broglie’s we can now use the first of de Broglie’s equation, eq. (1.12), we can replace $k = \frac{p}{\hbar}$ to obtain:

$$\frac{\partial^2\psi}{\partial x^2} = -\frac{p^2\psi}{\hbar^2},\tag{3.3}$$

which can be rearranged to:

$$p^2\psi = -\hbar^2 \frac{\partial^2\psi}{\partial x^2}.\tag{3.4}$$

The total energy associated with a wave moving in space is simply the sum of its kinetic and potential energies:

$$E = \frac{p^2}{2m} + V(x),\tag{3.5}$$

from which we can obtain:

$$p^2 = 2m[E - V(x)],\tag{3.6}$$

which we can then replace into eq. (3.4) to obtain:

$$2m[E - V(x)]\psi = -\hbar^2 \frac{\partial^2\psi}{\partial x^2},\tag{3.7}$$

which can then be rearranged to the famous **time-independent Schrödinger equation (TISEq)**:

$$-\frac{\hbar^2}{2m} \frac{\partial^2\psi}{\partial x^2} + V(x)\psi = E\psi,\tag{3.8}$$

A two-body problem can also be treated by this equation if the mass m is replaced with a reduced mass $\mu = \frac{m_1 m_2}{m_1 + m_2}$.

3.2 The Time-Dependent Schrödinger Equation

Unfortunately, the analogy with the classical wave equation that allowed us to obtain the TISEq in the previous section cannot be extended to the time domain by considering the equation that involves the partial first derivative with respect to time. Schrödinger himself presented his time-independent equation first, and then went back and postulated the more general time-dependent equation. We are following here the same strategy and just give the time-independent variable as a postulate. The single-particle time-dependent Schrödinger equation is:

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x)\psi(x, t) \quad (3.9)$$

where $V \in \mathbb{R}^n$ represents the potential energy of the system. Obviously, the time-dependent equation can be used to derive the time-independent equation. If we write the wavefunction as a product of spatial and temporal terms, $\psi(x, t) = \psi(x)f(t)$, then equation (3.9) becomes:

$$\psi(x)i\hbar \frac{df(t)}{dt} = f(t) \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x), \quad (3.10)$$

which can be rearranged to:

$$\frac{i\hbar}{f(t)} \frac{df(t)}{dt} = \frac{1}{\psi(x)} \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x). \quad (3.11)$$

Since the left-hand side of eq. (3.11) is a function of t only and the right hand side is a function of x only, the two sides must equal a constant. If we tentatively designate this constant E (since the right-hand side clearly must have the dimensions of energy), then we extract two ordinary differential equations, namely:

$$\frac{1}{f(t)} \frac{df(t)}{dt} = -\frac{iE}{\hbar} \quad (3.12)$$

and:

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x)}{\partial x^2} + V(x)\psi(x) = E\psi(x). \quad (3.13)$$

The latter equation is the TISEq. The former equation is easily solved to yield

$$f(t) = e^{-iEt/\hbar} \quad (3.14)$$

The solutions of eq. (3.14), $f(t)$, are purely oscillatory, since $f(t)$ never changes in magnitude. Thus if:

$$\psi(x, t) = \psi(x) \exp\left(\frac{-iEt}{\hbar}\right), \quad (3.15)$$

then the total wave function $\psi(x, t)$ differs from $\psi(x)$ only by a phase factor of constant magnitude. There are some interesting consequences of this. First of all, the quantity $|\psi(x, t)|^2$ is time independent, as we can easily show:

$$|\psi(x, t)|^2 = \psi^*(x, t)\psi(x, t) = \psi^*(x) \exp\left(\frac{iEt}{\hbar}\right) \psi(x) \exp\left(\frac{-iEt}{\hbar}\right) = \psi^*(x)\psi(x). \quad (3.16)$$

Wave functions of the form of eq. (3.15) are called stationary states. The state $\psi(x, t)$ is “stationary,” but the particle it describes is not! Of course eq. (3.14) represents only a particular solution to the time-dependent Schrödinger equation. The general solution is much more complicated, and the factorization of the temporal part is often not possible. The total wave function, however, can in general be written as a linear combination (superposition) of simpler terms, such as:¹

$$\psi(\mathbf{r}, t) = \sum_i c_i e^{-iE_i t/\hbar} \psi_i(\mathbf{r}). \quad (3.17)$$

3.3 Solution for the Free Particle

We will now show how to solve the time independent portion for the simple system of a particle in motion and not subject to any force. By definition, this free particle does not feel any external force, therefore $V(x) = 0$, and the TISEq is written simply:

¹This section was adapted in part from Prof. C. David Sherrill's A Brief Review of Elementary Quantum Chemistry Notes available here.

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi(x). \quad (3.18)$$

This equation can be rearranged to:

$$\frac{d^2\psi}{dx^2} = -\frac{2mE}{\hbar^2}\psi(x), \quad (3.19)$$

which corresponds to a mathematical problem where the second derivative of a function should be equal to a constant, $-\frac{2mE}{\hbar^2}$ multiplied by the function itself. The most general solution to this problem is the function:

$$\psi(x) = A \exp(ikx) + B \exp(-ikx), \quad (3.20)$$

with A and B arbitrary (complex) constants. The first and second derivatives of this function are:

$$\begin{aligned} \frac{d\psi(x)}{dx} &= ikA \exp(ikx) - ikB \exp(-ikx) \\ \frac{d^2\psi(x)}{dx^2} &= -k^2 A \exp(ikx) - k^2 B \exp(-ikx) = -k^2 \psi(x). \end{aligned} \quad (3.21)$$

Comparing the second derivative in eq. (3.21) with eq. (3.19), we immediately see that if we set:

$$k^2 = \frac{2mE}{\hbar^2}, \quad (3.22)$$

we solve the original differential equation. Considering de Broglie's equation, eq. (1.12), we can replace $k = \frac{p}{\hbar}$, to obtain:

$$E = \frac{k^2 \hbar^2}{2m} = \frac{p^2}{2m}, \quad (3.23)$$

which is exactly the classical value of the kinetic energy of a free particle moving in one direction of space. Since the function in eq. (3.20) solves the Schrödinger equation for the free particle, it is called an *eigenfunction* (or *eigenstate*) of the TISEq. The energy result of eq. (3.23) is called *eigenvalue* of the TISEq. Notice that, since k is continuous in the eigenfunction, the energy eigenvalue is also continuous (i.e., all values of E are acceptable).

3.4 Chapter Review

3.4.1 Study Questions

1. The starting point for deriving the 1D time-independent Schrödinger equation in Chapter 3 is a classical wave of which form?

- a. $\psi(x, t) = \exp[-kx + \omega t]$
 - b. $\psi(x, t) = \cos(kx - \omega t)$
 - c. $\psi(x, t) = \sin(kx + \omega t)$
 - d. $\psi(x, t) = \exp[i(kx - \omega t)]$
 - e. $\psi(x, t) = kx + \omega t$
-

2. For a two-body problem, the single-particle time independent Schrödinger equation can be used if the mass m is replaced by what?

- a. The total mass $m_1 + m_2$
 - b. The reduced mass $\mu = \frac{m_1 m_2}{m_1 + m_2}$
 - c. The mass difference $|m_1 - m_2|$
 - d. The geometric mean $\sqrt{m_1 m_2}$
 - e. Twice the lighter mass
-

3. Which of the following equation is the correct time-dependent Schrödinger equation for one particle?

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

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

d. $i\hbar \frac{\partial^2 \psi(x, t)}{\partial t^2} = V(x)\psi(x, t)$

e. $\frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2}$

4. For which of the following wave functions we can separate the variables in the Schrödinger equation?

a. $\psi(x, t) = \psi(x) + f(t)$

b. $\psi(x, t) = \psi(x)f(t)$

c. $\psi(x, t) = \psi(x)^2/f(t)$

d. $\psi(x, t) = xf(t)$

e. $\psi(x, t) = \psi(x)f(t)^2$

5. After separation of variables, the time-dependent Schrödinger equation yields two ordinary differential equations with a common constant. Which one?

a. The mass m

b. The momentum p

c. The angular frequency ω

d. The energy E

e. The wave number k

6. What is the time dependence of the probability density $|\psi(x, t)|^2$ of a stationary state?

- a. It increases linearly with time
 - b. It decreases exponentially with time
 - c. It oscillates sinusoidally with time
 - d. It is time independent
 - e. It diverges for large times
-

7. For the free particle in 1D, which of the following properties does the potential $V(x)$ have?

- a. It is a constant nonzero value
 - b. It is infinite at the origin
 - c. It is zero everywhere
 - d. It is linear in x
 - e. It is harmonic in x^2
-

8. Which of the following is a suitable solution for the free-particle time independent Schrödinger equation?

- a. $\psi(x) = Ae^{\pm ikx}$
- b. $\psi(x) = Ax$
- c. $\psi(x) = Ae^{kx}$
- d. $\psi(x) = A \cosh(kx)$

- e. $\psi(x) = A\delta(x)$
-

9. What is the energy eigenvalue for the free particle?

- a. $E = p^2$
 - b. $E = \frac{1}{2}mp^2$
 - c. $E = \hbar p$
 - d. $E = \frac{\hbar^2}{2mp^2}$
 - e. $E = \frac{p^2}{2m}$
-

10. What is a key difference between the free particle and bound systems regarding energy eigenvalues?

- a. Free-particle energies are always zero
 - b. Free-particle energies are always negative
 - c. Free-particle energies are always degenerate
 - d. Free-particle energies form a continuous spectrum, while bound systems have discrete levels
 - e. Free-particle energies do not depend on mass
-

Answers: 1.d, 2.b, 3.a, 4.b, 5.d, 6.d, 7.c, 8.a, 9.e, 10.d

Chapter 4

Analytically Solvable Models

The TISEq can be solved analytically only in a few special cases. We saw one of those special cases, the free particle, at the end of the previous chapter. In this section, we will analyze four others. Luckily, we can use these solutions to explain most of the effects in chemistry since we can combine them to describe the hydrogen atom upon which we can build more complex chemical systems, as we will show in the next chapters.

4.1 The Particle in a Box

We can start by considering a particle constrained to move in a single dimension, under the influence of a potential $V(x)$ which is zero for $0 \leq x \leq a$ and infinite elsewhere. Since the wavefunction is not allowed to become infinite, it must have a value of zero where $V(x)$ is infinite, so $\psi(x)$ is nonzero only within $[0, a]$. The Schrödinger equation is thus:

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi(x) \quad 0 \leq x \leq a. \quad (4.1)$$

In other words, inside the box $\psi(x)$ describes a free particle, but outside the box $\psi(x) = 0$. Since the Schrödinger equation involves derivatives, the function that solves it, $\psi(x)$, must be everywhere continuous and everywhere continuously differentiable. This fact means that the value of the wave function at the two extremes must be equal to zero:

$$\psi(0) = \psi(a) = 0. \quad (4.2)$$

Inside the box the solution will be the same as for the free particle:

$$\psi(x) = Ae^{ikx} + Be^{-ikx}, \quad (4.3)$$

where A and B are constants. Using the two constraints in eq. (4.2), we get at $x = 0$:

$$\psi(0) = A + B = 0 \implies B = -A \quad (4.4)$$

And at $x = a$:

$$\psi(a) = Ae^{ika} + Be^{-ika} = 0. \quad (4.5)$$

Replacing (4.4) into (4.5) gives:

$$Ae^{ika} - Ae^{-ika} = 0 \quad (4.6)$$

which is then easily simplified into:

$$e^{ika} = e^{-ika} e^{i2ka} = 1. \quad (4.7)$$

This equation is solved when:

$$2ka = 2n\pi, \quad (4.8)$$

which is trivially solved by either $k = 0$ or $a = 0$. This trivial solution is completely uninteresting, since it describes no particles in no boxes. A more interesting set of solutions however, are those for which:

$$k_n = \frac{n\pi}{a} \quad (n = 1, 2, 3, \dots, \infty), \quad (4.9)$$

which represents an infinite set of functions, $\psi_n(x)$, determined by a positive integer number n , called *quantum number*. Since these functions solve the TISEq, they are also called *eigenfunctions*, but they are not a continuous set, unlike in the previous case.

Using (4.4), the wave functions becomes:

$$\psi_n(x) = A \left(e^{i \frac{n\pi x}{a}} - e^{-i \frac{n\pi x}{a}} \right), \quad (4.10)$$

and using Euler's formula $\left(\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} \right)$, we get:

$$\psi_n(x) = A \frac{e^{i \frac{n\pi x}{a}} - e^{-i \frac{n\pi x}{a}}}{2i} = C \sin \left(\frac{n\pi x}{a} \right) \quad (4.11)$$

where $C = \frac{A}{i}$ is a complex constant.

To calculate the energy eigenvalues, we can replace (4.6) into eq. (4.1), to obtain:

$$E_n = n^2 \frac{\pi^2 \hbar^2}{2ma^2} = \quad (n = 1, 2, \dots, \infty). \quad (4.12)$$

A few interesting considerations can be made from the results of eq. (4.7). First, although there is an infinite number of acceptable values of the energy (eigenvalues), these values are not continuous. Second, the lowest value of the energy is not zero, and it depends on the size of the box, a , since:

$$E_1 = \frac{\pi^2 \hbar^2}{2ma^2} \neq 0. \quad (4.13)$$

This value is called *zero-point energy (ZPE)*, and is a purely quantum mechanical effect. Notice that we did not solve for the constant A . This task is not straightforward, and it can be achieved by requiring the wave function to describe one particle exclusively (we will come back to this task after chapter 7).

The most profound physical insight from the particle in a box solution comes from Born's interpretation of the wave function. The probability of finding the particle between positions x and $x + dx$ is given by $|\psi_n(x)|^2 dx$, where $|\psi_n(x)|^2$ is called the **probability density**.

For the particle in a box eigenfunctions in eq. (4.6), the probability density is:

$$|\psi_n(x)|^2 = A^2 \sin^2 \left(\frac{n\pi x}{a} \right). \quad (4.14)$$

This function tells us where the particle is *most likely* to be found if we measure its position.

4.1.1 Wave function analysis

The main results of the solution of the TISEq for the particle in a box are reported in figure 4.1. In this picture, the first five eigenfunctions are plotted together with their probability densities and spaced according to their energy levels. The non-trivial nodes are also reported.

A few important considerations can be made by analyzing the plots of the solutions, specifically:

1. **Eigenfunctions:** The solution of the TISEq for the particle in a box are mathematically identical to the **standing wave** modes on a vibrating string fixed at both ends. In classical waves, boundary conditions quantize allowed frequencies; in quantum mechanics, the infinite walls quantize energy via de Broglie waves, yielding stationary states that don't propagate but "stand" with fixed nodes.
2. **Eigenvalues:** The energy levels are not equidistant from each other, with the gap increasing linearly with $2n+1$: low-lying levels are closer together, and higher levels get progressively farther apart.
3. **Nodes:** At $x = 0$, $x = a$, and at $x = \frac{ka}{n}$ (where $k = 1, 2, \dots, n-1$), we have $\sin\left(\frac{n\pi x}{a}\right) = 0$, so $|\psi_n(x)|^2 = 0$. These are **nodes** where the probability of finding the particle is exactly zero. The boundary nodes, however, are rather trivial, as they are forced by the infinite potential walls, and not a direct consequence of the wavefunction's oscillatory nature. The n -th energy eigenstate has exactly $(n-1)$ nodes *inside* the box.
4. **Peak probability:** The maxima of $|\psi_n(x)|^2$ occur where $\sin^2\left(\frac{n\pi x}{a}\right)$ is maximum. These **antinodes** represent regions of highest probability.

For the ground state ($n = 1$), $|\psi_1(x)|^2 = A^2 \sin^2\left(\frac{\pi x}{a}\right)$ has one maximum at $x = a/2$. The particle is most likely to be found at the center of the box. For the first excited state ($n = 2$), $|\psi_2(x)|^2 = A^2 \sin^2\left(\frac{2\pi x}{a}\right)$ has a node at $x = a/2$ and two equal maxima at $x = a/4$ and $x = 3a/4$. The particle has zero probability of being at the center.

Unlike classical mechanics where a particle with definite energy bounces back and forth uniformly across the box, quantum stationary states have non-uniform position probability distributions. The particle doesn't "spend equal time everywhere" - instead, it's more

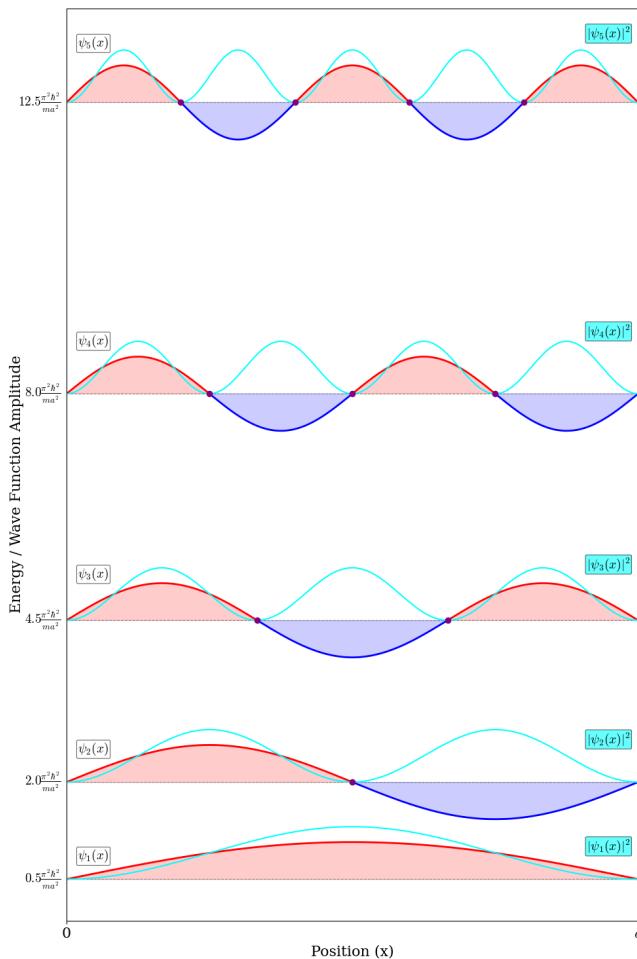


Figure 4.1: The first five eigenfunctions and energy levels of a particle confined inside a box by walls with infinite potential energy. The nodes of the wave functions are reported as purple dots. The probability density for each level is also plotted in cyan.

likely to be found near the antinodes of the wave function. Notice that we can use the trigonometric identity, $\sin^2 \theta = \frac{1-\cos(2\theta)}{2}$, to obtain:

$$|\psi_n(x)|^2 = A^2 \left[1 - \cos \left(\frac{2n\pi x}{a} \right) \right]. \quad (4.15)$$

This shows that the probability density is the sum of a uniform background (A^2) modulated by a rapidly oscillating term whose frequency increases with n . As n becomes large (high energy states), $|\psi_n(x)|^2$ becomes more uniform across the box, approaching the classical limit.

The connection between nodes, energy, and probability distribution is fundamental. Higher energy states oscillate more rapidly and have more nodes, reflecting their increased “kinetic energy” through faster spatial variation.

4.2 The Particle on a Ring

We can extend our understanding of confined quantum systems by considering a particle constrained to move in a circle of fixed radius r . This model, known as the particle on a ring (or particle in a ring), represents the rotational motion of a particle about a fixed axis.

While this system seems to be a 2-dimensional problem, in reality it is exactly the same problem as the particle in a box, with the two endpoints of the box wrapping around themselves and touching each other. However, rather than working with a single linear coordinate, x , we can use the set of polar coordinates, (r, ϕ) , where only the azimuthal angle ϕ varies (with $0 \leq \phi \leq 2\pi$). The radial coordinate r is held constant. The wave function $\psi(x)$ can be easily written as $\psi(\phi)$ by using $\phi = x/r$. Thus the second derivative becomes:

$$\frac{d^2\psi(x)}{dx^2} = \frac{1}{r^2} \frac{d^2\psi(\phi)}{d\phi^2}, \quad (4.16)$$

and the time-independent Schrödinger equation can be obtained directly from eq. (4.1) as:

$$-\frac{\hbar^2}{2mr^2} \frac{d^2\psi(\phi)}{d\phi^2} = E\psi(\phi). \quad (4.17)$$

For a classical particle rotating about the origin in the xy -plane, its kinetic energy is expressed as:

$$E = \frac{L_z^2}{2I}, \quad (4.18)$$

where $I = mr^2$ is the moment of inertia of the particle about the axis of rotation and L_z is the component of the angular momentum along the z axis, perpendicular to the rotation plane. The TISEq can then be rewritten as:

$$-\frac{\hbar^2}{2I} \frac{d^2\psi(\phi)}{d\phi^2} = E\psi(\phi), \quad (4.19)$$

where it is now evident that the energy of the quantum mechanical system is directly related to the component L_z of the angular momentum, as in the classical case. Since the angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, \mathbf{L} must point in the z -direction when \mathbf{r} and \mathbf{p} both lie on the xy -plane.

The boundary condition for the particle in a ring comes from the requirement that the wavefunction must be single-valued, meaning that as we traverse the ring from angle $\phi = 0$ to $\phi = 2\pi$, we must return to the same value:

$$\psi(\phi) = \psi(\phi + 2\pi) \quad (4.20)$$

Similarly to the previous cases, the general solution to the TISEq above is:

$$\psi(\phi) = Ae^{ik\phi} + Be^{-ik\phi} \quad (4.21)$$

where A and B are constants. Applying the single-valuedness boundary condition requires that $e^{ik \cdot 2\pi} = 1$, which is satisfied only when $k = m_\ell$, where m_ℓ is an integer (positive, negative, or zero) called *azimuthal quantum number*. The set of acceptable solutions is therefore:

$$\psi_{m_\ell}(\phi) = Ae^{im_\ell\phi} \quad (m_\ell = 0, \pm 1, \pm 2, \dots, \pm \infty). \quad (4.22)$$

The $m_\ell = 0$ solution is equivalent to the $n = 1$ solution of the particle in a box, while the remaining solutions represent vibrations in the clockwise (positive solutions, L_z pointing upwards) and counter-clockwise (negative solutions, L_z pointing downwards) directions.

To find the energy eigenvalues, we substitute the eigenfunction back into the Schrödinger equation. Taking the second derivative of $\psi_{m_\ell}(\phi)$:

$$\frac{d^2\psi_{m_\ell}}{d\phi^2} = -m_\ell^2 e^{im_\ell\phi} = -m_\ell^2 \psi_{m_\ell}(\phi) \quad (4.23)$$

Replacing this derivative in the Schrödinger equation, we obtain:

$$-\frac{\hbar^2}{2mr^2}(-m_\ell^2)\psi_{m_\ell} = E\psi_{m_\ell}, \quad (4.24)$$

and the final energy levels, as:

$$E_{m_\ell} = \frac{\hbar^2 m_\ell^2}{2mr^2} = \frac{\hbar^2 m_\ell^2}{2I}. \quad (4.25)$$

The energy levels of the particle on a ring are plotted for the first five levels in figure 4.2.

Using the definition of kinetic energy for a classical system in eq. (4.18), we get:

$$\frac{L_z^2}{2I} = \frac{\hbar^2 m_\ell^2}{2I}, \quad (4.26)$$

which implies that the angular momentum along the z -axis is quantized in units of \hbar as:

$$L_z = \hbar m_\ell \quad (m_\ell = 0, \pm 1, \pm 2, \dots, \pm \infty). \quad (4.27)$$

A rotating charged particle orbiting the ring generates a magnetic moment $\mu \propto \mathbf{L}$. Thus eq. (4.27), which quantizes L_z , explains why m_ℓ is called the **magnetic quantum number**. Several important observations can be made from these results. First, the energy eigenvalues form a discrete set, as expected for a confined system. Second, except for the ground state ($m_\ell = 0$), each energy level is doubly degenerate: the states with m_ℓ and $-m_\ell$ have the same energy. This makes physical sense because rotation in the positive and negative directions (clockwise and counter-clockwise) should have the same energy. Third, the energy does not depend on the sign of m_ℓ , which is a consequence of the fact that energy depends on L_z^2 , and $(-m_\ell)^2 = (m_\ell)^2$. This is the first instance we have seen of a degeneracy arising from the symmetry of the system itself (rotational symmetry).

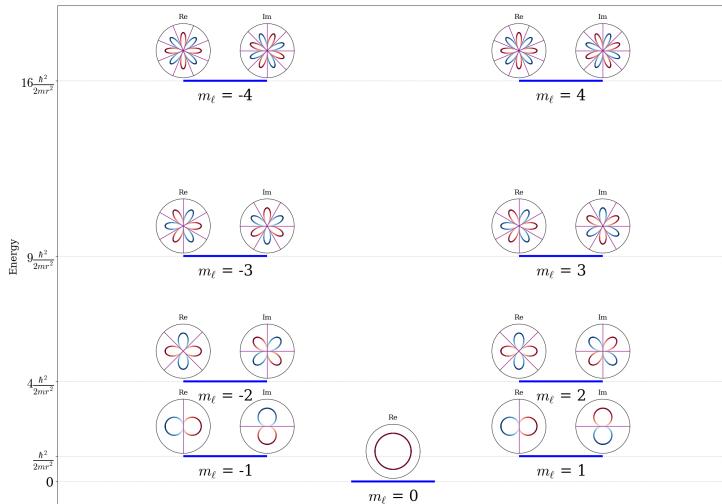


Figure 4.2: Energy level diagram and degeneracy pattern for the particle-on-a-ring. Both the real and imaginary parts of the wave functions of the first five energy levels are represented using polar plots (see text for explanation), with nodes reported as purple diameters of the physical ring (black circle).

4.2.1 Wave function analysis

The eigenfunctions in eq. (4.23) are complex functions and their visualization is not as simple as for the particle in a box. For $m_\ell \neq 0$, these represent states with definite angular momentum $L_z = \hbar m_\ell$. The probability density is:

$$|\psi_{m_\ell}(\phi)|^2 = |A|^2, \quad (4.28)$$

which is **uniform around the ring**. This means that the particle is equally likely to be found at any point on the ring in these states, consistent with our expectation that a state with definite angular momentum should have no preferred position on the ring.

Visualizing the particle-on-a-ring wave functions presents a challenge: the system is **intrinsically 1D** (motion along the circle), yet the **ring geometry requires 2D projection** for intuitive understanding. The clearest visualization plots the real and imaginary parts of $\psi_{m_\ell}(\phi)$ as height above/below the ring in the xy -plane, as in figure 4.3:

- xy -plane: Shows the circular path ($\phi = 0$ to 2π)
- z -axis: Real part $\Re[\psi_{m_\ell}] = \frac{1}{\sqrt{2\pi}} \cos(m_\ell \phi)$ (left column)
- z -axis: Imaginary part $\Im[\psi_{m_\ell}] = \frac{1}{\sqrt{2\pi}} \sin(m_\ell \phi)$ (right column)

This 3D representation preserves both the physical ring constraint and reveals the oscillatory nodal structure as ϕ varies around the circle.

Alternatively, **2D contour plots** map the real and imaginary parts of $\psi_{m_\ell}(\phi)$ onto the ring using a color gradient, as it's done in figure 4.4:

- Red shades (darker = more positive amplitude).
- Blue shades (darker = more negative amplitude).
- White/light colors = near zero (nodes).

These are top-down views of the 3D plots in figure 4.3, excellent for identifying nodal positions. In the particle-in-a-box, nodes were single points along a line. For the particle-on-a-ring, these clots show that nodes become pairs of diametrically opposite points on the circle. The diametric pairing reflects the rotational symmetry: if $\cos(m_\ell \phi_0) = 0$, then $\cos(m_\ell(\phi_0 + \pi)) = -\cos(m_\ell \phi_0) = 0$.

genfunctions can also be visualized using **polar plots** against the azimuthal angle ϕ , as shown in figure 4.5:

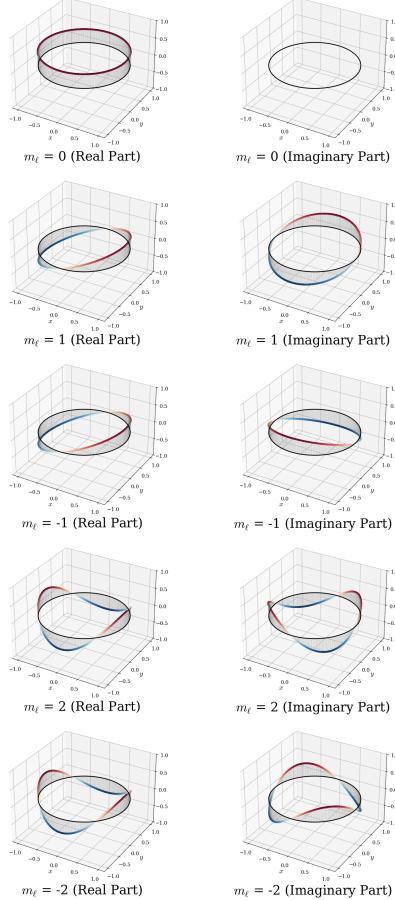


Figure 4.3: 3D visualization of particle-on-a-ring eigenfunctions $\psi_{m_\ell}(\phi) = \frac{1}{\sqrt{2\pi}} e^{im_\ell \phi}$. Left column: Real part $\Re[\psi_{m_\ell}] = \frac{1}{\sqrt{2\pi}} \cos(m_\ell \phi)$. Right column: Imaginary part $\Im[\psi_{m_\ell}] = \frac{1}{\sqrt{2\pi}} \sin(m_\ell \phi)$. Each row shows states with quantum numbers $m_\ell = 0, +1, -1, +2, -2$. The black circle represents the ring of fixed radius r . The colored surface shows wavefunction amplitude plotted as height (z -axis) above/below the ring as the azimuthal angle ϕ varies from 0 to 2π . Each component has exactly $2|m_\ell|$ nodal points, with $\pm m_\ell$ pairs showing opposite phase winding directions (degenerate in energy).

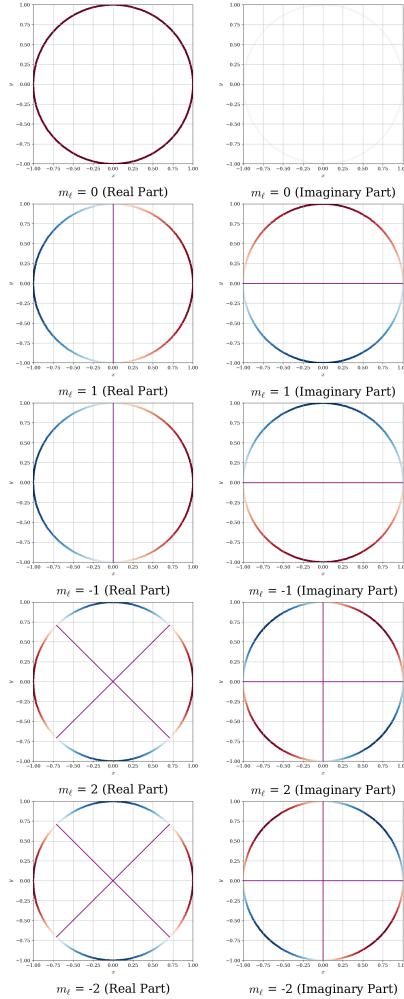


Figure 4.4: Top-down 2D contour plots of particle-on-a-ring eigenfunctions $\psi_{m_\ell}(\phi) = \frac{1}{\sqrt{2\pi}} e^{im_\ell \phi}$. Left column: Real part $\text{Re}[\psi_{m_\ell}] = \frac{1}{\sqrt{2\pi}} \cos(m_\ell \phi)$. Right column: Imaginary part $\text{Im}[\psi_{m_\ell}] = \frac{1}{\sqrt{2\pi}} \sin(m_\ell \phi)$. Each row shows states with quantum numbers $m_\ell = 0, +1, -1, +2, -2$. Color scheme: Dark red (positive amplitude), dark blue (negative amplitude), white/light colors (nodes ≈ 0). Each component shows exactly $2|m_\ell|$ nodal points appearing as diametrically opposite pairs due to rotational symmetry.

- The **radial distance** from the center represents the **amplitude** of the real or imaginary part of the wavefunction.
- The **angular position** corresponds to **azimuthal angle** (ϕ).

This representation clearly reveals the nodes where amplitude = 0, appearing as returns to the origin, and the lobes, that is regions of maximum amplitude. However, this visualization has a limitation, as it completely loses the physical ring geometry. The particle remains confined to a circle of fixed radius r , but the polar plot makes it appear as if it explores a 2D disk. When observing the polar plots, it is important to keep in mind that **the lobes are not spatial probability distributions** in 2D space. They represent the **magnitude of the real/imaginary parts** of $\psi_{m_\ell}(\phi)$ evaluated along the ring. The actual position probability $|\psi|^2 = |A|^2$ remains perfectly uniform around the circle for all states.

4.3 The Particle on a Sphere

Extending the particle on a ring to a three-dimensional case, we get a particle confined on the surface of a sphere (particle on a sphere). This system is a 2-dimensional system analogous to the motion of a rigid rotor. The rigid rotor is a simple model of a rotating stick in three dimensions. We consider the stick to consist of two point-masses at a fixed distance. We then reduce the model to a single-particle system by considering the rigid rotor to have one mass fixed at the origin, which is orbited by the reduced mass μ , at a distance r .¹ The Cartesian coordinates, x, y, z , are not optimal to represent the system, and they are usually replaced by three spherical polar coordinates: the co-latitude (zenith) angle θ , the longitudinal (azimuth) angle ϕ , and the distance r . If the orbiting mass (the electron) is held at a constant distance from the origin, the system reduces to a 2-dimensional system by simply keeping r constant. The TISEq of the system in spherical coordinates is:

$$-\frac{\hbar^2}{2I} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] \psi(r) = E\psi(r), \quad (4.29)$$

where $I = \mu r^2$ is the moment of inertia, as for the particle in a ring case. After a little effort, the eigenfunctions can be shown to be the complex

¹Notice that if we consider an electron rotating at a fixed distance around a nucleus, then the difference in the mass of the electron and the reduced mass of the electron-nucleus system is to all intent and purposes negligible, and we can treat this case as a particle on a sphere.

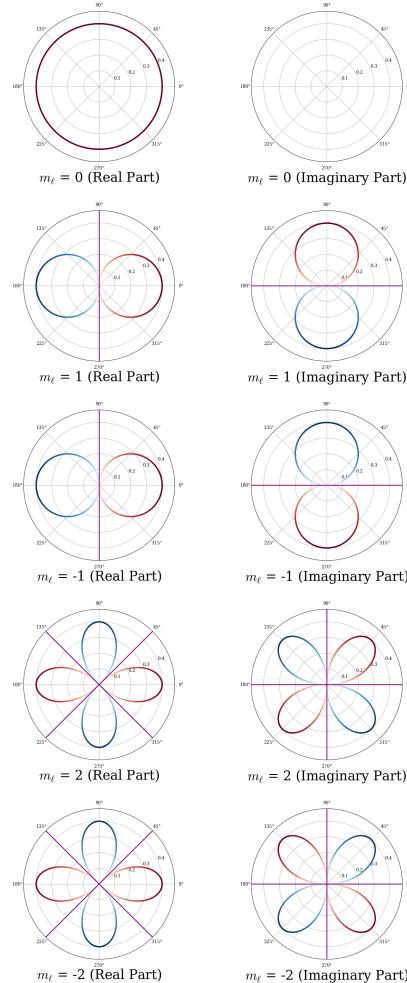


Figure 4.5: Polar plots of particle-on-a-ring eigenfunctions $\psi_{m_\ell}(\phi) = \frac{1}{\sqrt{2\pi}} e^{im_\ell \phi}$. Left column: Real part $\Re[\psi_{m_\ell}] = \frac{1}{\sqrt{2\pi}} \cos(m_\ell \phi)$. Right column: Imaginary part $\Im[\psi_{m_\ell}] = \frac{1}{\sqrt{2\pi}} \sin(m_\ell \phi)$. Each row shows states with quantum numbers $m_\ell = 0, +1, -1, +2, -2$. Radial distance from center \propto wavefunction amplitude; angular position $= \phi$. Nodes appear where the curve returns to the origin ($2|m_\ell|$ per component). Note: Lobes represent amplitude along the ring, not 2D spatial probability (actual $|\psi|^2$ is uniform).

spherical harmonics:²

$$Y_\ell^{m_\ell}(\theta, \phi) = N P_\ell^{|m_\ell|}(\cos \theta) e^{im_\ell \phi}, \quad (4.30)$$

where $P_\ell^{|m_\ell|}(\cos \theta)$ denotes the associated Legendre polynomials, m_ℓ is the **magnetic quantum number**, and N is a normalization constant. As for the previous cases, the eigenvalues are quantized and are simply:

$$E_\ell = \frac{\hbar^2}{2I} \ell(\ell + 1) \quad (\ell = 0, 1, 2, \dots, \infty). \quad (4.31)$$

The spherical harmonic functions depend on a second quantum number, $m_\ell = -\ell, -\ell+1, \dots, \ell-1, \ell$ which is the same **magnetic quantum number** that we encounter for the particle on a ring. For the spherical harmonics, however, the energy does not depend on the magnetic quantum number, and functions with the same ℓ but different m_ℓ are energy-degenerate. Each energy level E_ℓ is $(2\ell + 1)$ -fold degenerate in m_ℓ . Surprisingly enough, this problem was, after all, a one-dimensional problem in disguise, similarly to the previous two cases. The two angular variables, (θ, ϕ) , that enters the wave function are indeed completely identical to each other, with one of the two determining the energy (the magnitude of the angular momentum of the circular motion) and the other determining only an energetically irrelevant orientation for the circular motion. Notice, however, that both these components are quantized. So, for a rigid rotor, the magnitude of the angular momentum along the z -axis is quantized in units of \hbar and jumps discontinuously from one value to the next, as for the particle on a ring case.³ For the particle-on-a-sphere case, however, the orientation in space at which the rotor is allowed to rotate is also quantized and it changes discontinuously. The energy levels of the particle on a sphere are plotted for the first five levels in figure 4.6.

4.3.1 Wave function analysis

The analytical expressions for the first three energy levels ($\ell = 0, 1, 2$) of the rigid rotor are:

$\ell = 0$:

²For a description of the spherical harmonics see here

³You can think of this as either the “speed” at which the rotor is rotating, or even better, as its “angular frequency”.



Figure 4.6: Energy level diagram and degeneracy pattern for the particle-on-a-sphere.

$$Y_0^0(\theta, \phi) = \frac{1}{2} \sqrt{\frac{1}{\pi}}.$$

$\ell = 1$:

$$Y_1^{-1}(\theta, \phi) = \frac{1}{2} \sqrt{\frac{3}{2\pi}} e^{-i\phi} \sin \theta = \frac{1}{2} \sqrt{\frac{3}{2\pi}} \frac{x-iy}{r}.$$

$$Y_1^0(\theta, \phi) = \frac{1}{2} \sqrt{\frac{3}{\pi}} \cos \theta = \frac{1}{2} \sqrt{\frac{3}{\pi}} \frac{z}{r}.$$

$$Y_1^1(\theta, \phi) = -\frac{1}{2} \sqrt{\frac{3}{2\pi}} e^{i\phi} \sin \theta = -\frac{1}{2} \sqrt{\frac{3}{2\pi}} \frac{x+iy}{r}.$$

$\ell = 2$:

$$Y_2^{-2}(\theta, \phi) = \frac{1}{4} \sqrt{\frac{15}{2\pi}} e^{-2i\phi} \sin^2 \theta = \frac{1}{4} \sqrt{\frac{15}{2\pi}} \frac{(x-iy)^2}{r^2}.$$

$$Y_2^{-1}(\theta, \phi) = \frac{1}{2} \sqrt{\frac{15}{2\pi}} e^{-i\phi} \sin \theta \cos \theta = \frac{1}{2} \sqrt{\frac{15}{2\pi}} \frac{(x-iy)z}{r^2}.$$

$$Y_2^0(\theta, \phi) = \frac{1}{4} \sqrt{\frac{5}{\pi}} (3 \cos^2 \theta - 1) = \frac{1}{4} \sqrt{\frac{5}{\pi}} \frac{3z^2 - r^2}{r^2}.$$

$$Y_2^1(\theta, \phi) = -\frac{1}{2} \sqrt{\frac{15}{2\pi}} e^{i\phi} \sin \theta \cos \theta = -\frac{1}{2} \sqrt{\frac{15}{2\pi}} \frac{(x+iy)z}{r^2}.$$

$$Y_2^2(\theta, \phi) = \frac{1}{4} \sqrt{\frac{15}{2\pi}} e^{2i\phi} \sin^2 \theta = \frac{1}{4} \sqrt{\frac{15}{2\pi}} \frac{(x+iy)^2}{r^2}.$$

Visualizing spherical harmonics presents the same challenges as the particle on a ring, but amplified by dimensionality. The particle lives on a 2D spherical surface (requiring 3D embedding), and we need to show both position (θ, ϕ) and amplitude $(Y_\ell^{m_\ell})$ —a 4D problem squeezed into 3D plots.

The solution mirrors the ring contour plots: we map the amplitude directly to a color gradient on the unit sphere surface. Red shades represent positive values of $\Re[Y_\ell^{m_\ell}(\theta, \phi)]$ or $\Im[Y_\ell^{m_\ell}(\theta, \phi)]$, blue shades represent negative values, and white/light colors mark the nodes where the harmonic equals zero. Figure 4.7 shows these color maps for the real parts across $\ell = 0, 1, 2$, while figure 4.8 shows the corresponding imaginary parts.

This approach preserves the spherical geometry while clearly revealing the nodal lines as white curves separating red and blue regions, and makes the lobe structure immediately visible—directly extending the 1D ring visualization strategy to the 2D sphere.

Crucially, these plots visualize the wavefunction itself, not the probability density. The actual position probability on the sphere is given by $|Y_\ell^{m_\ell}(\theta, \phi)|^2$. Unlike the particle on a ring case, where $|\psi_{m_\ell}(\phi)|^2 = |A|^2$ was perfectly uniform, the spherical harmonics have non-uniform angular probability. For example:

- Y_0^0 : Uniform probability everywhere on sphere

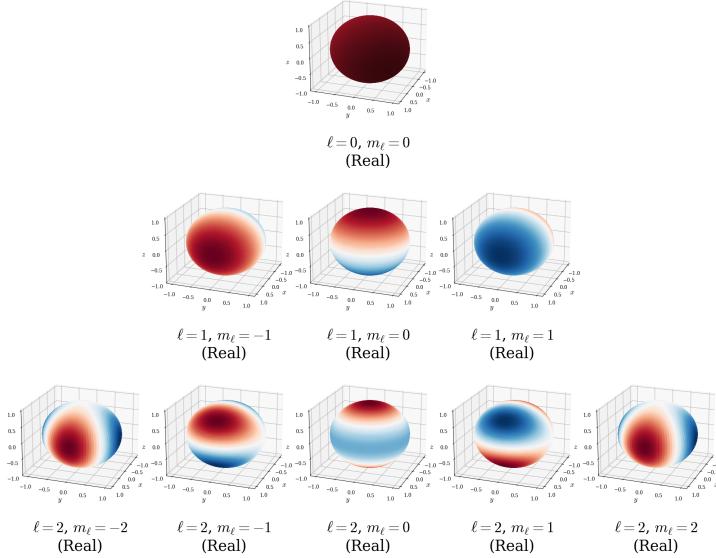


Figure 4.7: Color gradient maps of the real part ($\Re[Y_\ell^{m_\ell}(\theta, \phi)]$) for spherical harmonics of the rigid rotor. Color scheme: Red (positive), blue (negative), white/light (nodal graticules where $Y_\ell^{m_\ell} = 0$). Each ℓ manifold shows characteristic nodal structure: 0 nodal lines ($\ell = 0$), 1 nodal surface ($\ell = 1$), 2 nodal surfaces ($\ell = 2$).

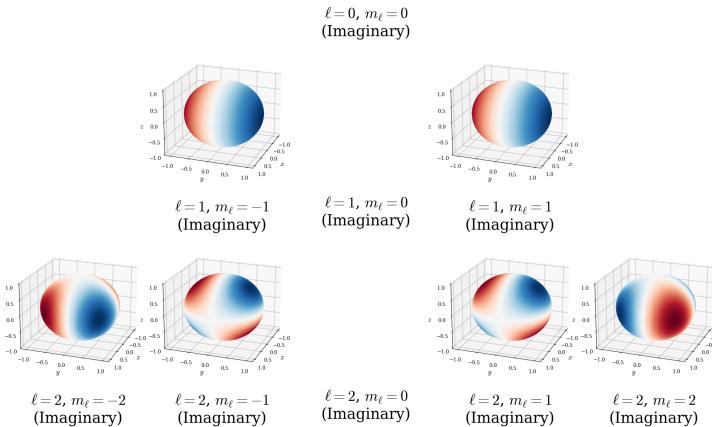


Figure 4.8: Color gradient maps of the imaginary part ($\Im[Y_\ell^{m\ell}(\theta, \phi)]$) for spherical harmonics of the rigid rotor. Color scheme: Red (positive), blue (negative), white/light (nodal lines where $Y_\ell^{m\ell} = 0$). Each ℓ manifold shows characteristic nodal structure: 0 nodal lines ($\ell = 0$), 1 nodal surface ($\ell = 1$), 2 nodal surfaces ($\ell = 2$).

- $Y_1^0 \propto \cos \theta$: Probability peaks at poles, zero at equator
- $Y_1^{\pm 1} \propto \sin \theta$: Probability peaks at equator, zero at poles
- Y_2^0 : Probability concentrated along z-axis with equatorial minimum

To visualize the true probability distribution, we can plot $|Y_\ell^{m_\ell}|^2$ directly, which eliminates phase information (\Re, \Im) but reveals where the particle is most likely to be found on the sphere's surface, as in figure 4.9. The loss of phase is clear from the fact that the positive and negative m_ℓ functions on these plots look exactly identical. A simple interpretation of this fact is that the phase is associated with direction of rotation: Functions with positive m_ℓ represent clockwise rotation (L_z pointing upwards), while functions with negative m_ℓ represent counter-clockwise rotation (L_z pointing downwards), as for the particle on a ring.

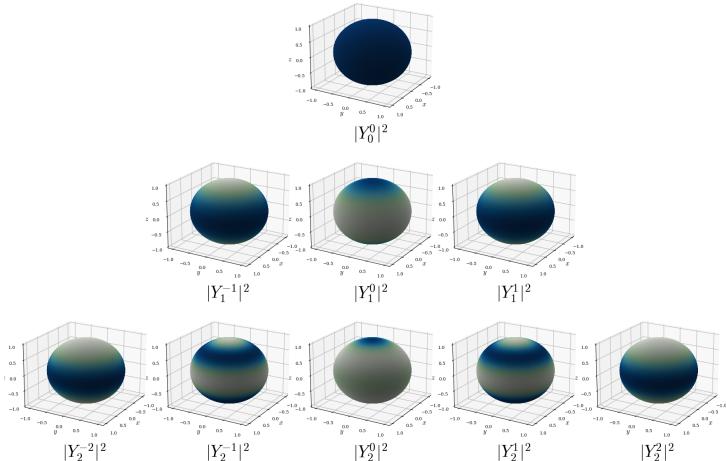


Figure 4.9: Probability density plots $|Y_\ell^{m_\ell}(\theta, \phi)|^2$ for spherical harmonics of the rigid rotor. Color scheme: Darker colors indicate higher probability density on the surface of the sphere. Unlike the particle on a ring ($|\psi|^2$), these show non-uniform angular probability: ($\ell = 0$) uniform; ($\ell = 1$) states peak at poles/equator; ($\ell = 2$) states show complex 4-lobed or toroidal patterns. Probability integrates to 1 over sphere surface for each state.

An alternative, and often more intuitive, representation is obtained by plotting the magnitude of the spherical harmonic as a radial deformation of the sphere. In this parametric view, each point on the unit sphere (θ, ϕ) is moved

radially outward (or inward) by an amount proportional to $|Y_\ell^{m_\ell}(\theta, \phi)|$. The deformed surface is then colored to indicate the sign of the real part (for example, red for positive, blue for negative). This results in a parametric plot where the distance from the center at a given angle represents the magnitude of the amplitude of the spherical harmonic. These plots are reported in figure 4.10 for the real part of the wave functions, and in figure 4.11 for the imaginary part. These deformed plots correspond to the polar plots for the particle on a ring case we reported in figure 4.4.

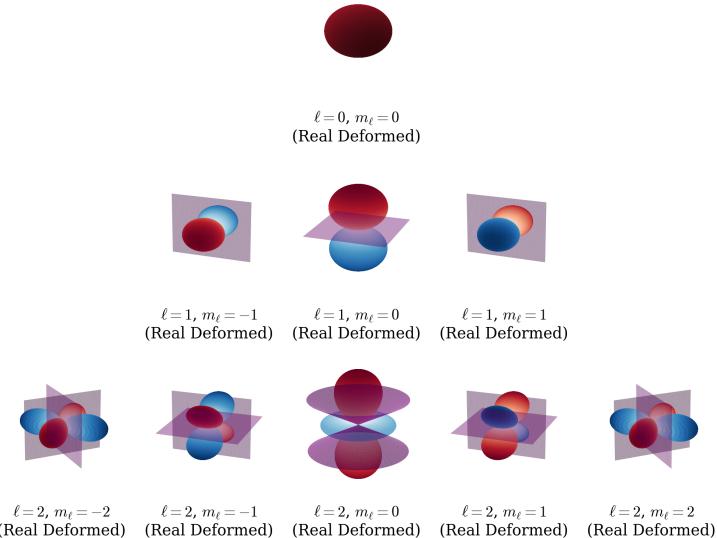


Figure 4.10: Parametric plots of the real part of the spherical harmonics $Y_\ell^{m_\ell}(\theta, \phi)$. The radial deformation (distance from the center at a given angle) is proportional to $\Re[Y_\ell^{m_\ell}(\theta, \phi)]$ (larger lobes = higher angular probability). The color indicates sign of real part (red=positive, blue=negative). The nodal surfaces are also reported in purple. Note: The particle remains confined to a constant sphere surface; the lobes visualize angular probability structure, not 3D spatial extent.

The parametric plots are ideal to show the **nodal structure** of the spherical harmonics. This nodal structure follows a simple rule: the ℓ -th energy level has exactly ℓ nodal surfaces on the sphere. These nodal surfaces—where $Y_\ell^{m_\ell} = 0$ —appear as *nodal circles* on the spherical surface, as in figures 4.7

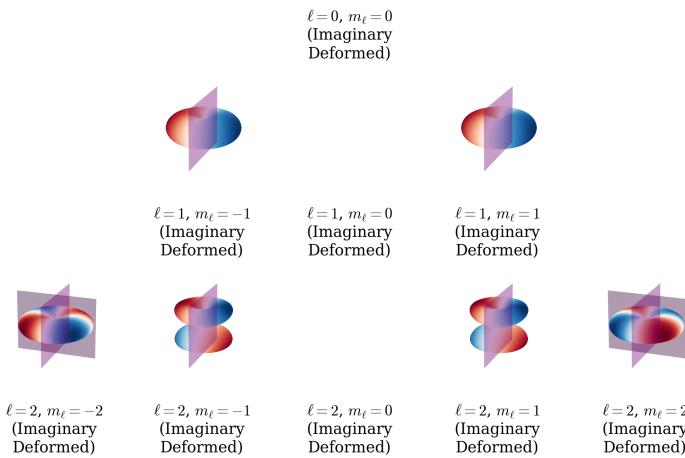


Figure 4.11: Parametric plots of the real part of the spherical harmonics $Y_\ell^{m_\ell}(\theta, \phi)$. The radial deformation (distance from the center at a given angle) is proportional to $\Im[Y_\ell^{m_\ell}(\theta, \phi)]$ (larger lobes = higher angular probability). The color indicates sign of real part (red=positive, blue=negative). The nodal surfaces are also reported in purple. Note: The particle remains confined to a constant sphere surface; the lobes visualize angular probability structure, not 3D spatial extent.

and 4.8. Most of these nodes are great circles on the sphere, but in some case they might not be, as they might be parallel nodes along the z -axis (see i.e. Y_2^0). In parametric plots (deformed polar views), the nodes appear as flat planes cutting through the origin when they are great circles, while they appear as cones when they are parallel circles along the z -axis.

Finally, it is extremely important to notice that although the deformed polar view resembles the standard “atomic orbital” pictures from general chemistry, **they are fundamentally different visualizations with distinct physical meanings**. Specifically, the deformed lobes do not represent spatial probability in 3D space. Just as with the particle on a ring, the particle remains confined to the sphere’s surface at fixed radius, r . The parametric deformation makes it appear as if the particle explores the 3D volume inside the sphere—**it does not**. When observing the polar plots, it is important to keep in mind that the lobes are not spatial probability distributions in 3D space, but they represent the magnitude of the real/imaginary parts of the spherical harmonic evaluated on the surface of the sphere. As we already discussed above, the probability is not uniform—unlike the particle-on-a-ring where $|\psi|^2 = |A|^2$ was constant. For the rigid rotor, $|Y_\ell^{m_\ell}|^2$ has maxima and minima (including nodes), showing the electron is more likely found in certain angular directions than others, as it’s clearly shown in figure 4.9. The deformed polar view visualizes this angular probability structure, but remains a 2D angular function plotted in a deformed 3D surface, not a 3D spatial probability.

4.4 The Harmonic Oscillator

We now consider a particle subject to a restoring force $F = -kx$, as might arise for a mass-spring system obeying Hooke’s Law. The potential is then:

$$V(x) = - \int_{-\infty}^{\infty} (-kx) dx = V_0 + \frac{1}{2} kx^2. \quad (4.32)$$

If we choose the energy scale such that $V_0 = 0$ then: $V(x) = \frac{1}{2} kx^2$, and the TISEq looks:

$$-\frac{\hbar^2}{2\mu} \frac{d^2\psi}{dx^2} + \frac{1}{2} kx^2 \psi(x) = E\psi(x) \quad (4.33)$$

After some effort, the eigenfunctions are:

$$\psi_n(x) = N_n H_n(\alpha^{1/2} x) e^{-\alpha x^2/2} \quad (n = 0, 1, 2, \dots, \infty), \quad (4.34)$$

where H_n is the Hermite polynomial of degree n , and α and N_n are defined by

$$\alpha = \sqrt{\frac{k\mu}{\hbar^2}} \quad N_n = \frac{1}{\sqrt{2^n(n)!}} \left(\frac{\alpha}{\pi} \right)^{1/4}. \quad (4.35)$$

The eigenvalues are:

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right), \quad (4.36)$$

with $\omega = \sqrt{k/\mu}$. Notice how, once again, the eigenfunctions and eigenvalues are not continuous, and the lowest energy state is not zero. In other words, the two masses of a quantum harmonic oscillator are always in motion. The frequencies at which they vibrate do not form a continuous spectrum. That is, the vibration frequency cannot take any value that we can think of, but only those given by eq. (4.36). The lowest possible energy (the ZPE) will be:

$$E_0 = \frac{1}{2}\hbar\omega \neq 0 \quad (4.37)$$

4.4.1 Wave function analysis

Figure 4.12 shows the first five harmonic oscillator eigenfunctions alongside their probability densities, with vertical positions scaled according to their energy levels. The nodal positions are also marked.

Similarly to the particle in a box case, each n -th eigenstate has exactly n interior nodes, with the Gaussian envelope $e^{-\xi^2/2}$ localizing probability near $x = 0$ and ensuring rapid decay at large $|x|$. The ground state ($n = 0$) is a simple Gaussian with no nodes. The first excited state ($n = 1$) has one node at $x = 0$, and higher states develop increasingly rapid oscillations within the narrowing classical turning points. In contrast with the particle in a box case, the energy levels are equally spaced, with constant spacing $\Delta E = \hbar\omega$, reflecting the classical periodicity of simple harmonic motion.

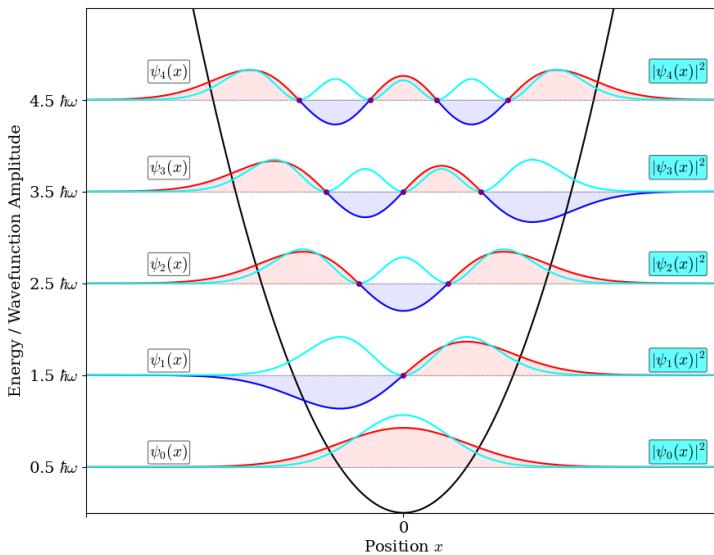


Figure 4.12: The first five eigenfunctions and energy levels of a particle confined inside a harmonic oscillator potential. The nodes of the wave functions are reported as purple dots. The probability density for each level is also plotted in cyan.

The probability densities $|\psi_n(x)|^2$ remain Gaussian-localized near $x = 0$ for all n . Higher- n states oscillate faster within progressively tighter classical turning points, with classical probability peaking near maximum displacement where the particle moves slowest.

4.5 Chapter Review

4.5.1 Study Questions

1. Which three model systems are treated in this chapter as analytically solvable?

- a. Hydrogen atom, particle in a box, rigid rotor
 - b. Hydrogen atom, harmonic oscillator, Morse oscillator
 - c. Particle in a box, harmonic oscillator, rigid rotor
 - d. Free particle, hydrogen atom, rigid rotor
 - e. Particle in a box, Coulomb potential, rigid rotor
-

2. For the 1D infinite potential well (“particle in a box”), the potential $V(x)$ is what?

- a. Constant for all x
 - b. Zero for $0 \leq x \leq a$, infinite elsewhere
 - c. Linear in x between 0 and a
 - d. Harmonic in x^2 between 0 and a
 - e. Zero everywhere
-

3. What boundary conditions are imposed on the wavefunction for the particle in a box?

- a. $\psi(0) = \psi(a) = 1$
 - b. $\psi(0) = 0, \psi(a) = 1$
 - c. $\psi(0) = \psi'(a) = 0$
 - d. $\psi'(0) = \psi'(a) = 0$
 - e. $\psi(0) = \psi(a) = 0$
-

4. Applying the $\psi(0) = 0$ boundary condition to the particle in a box leads to which condition on the constants??

- a. $A = 0$
 - b. $k = 0$
 - c. $A = B$
 - d. $B = 0$
 - e. $A = -B$
-

5. Which of the following is the correct spatial dependence of the normalized eigenfunctions for the particle in a box?

- a. $\psi_n(x) \propto \sin\left(\frac{n\pi x}{a}\right)$
- b. $\psi_n(x) \propto \cos\left(\frac{n\pi x}{a}\right)$
- c. $\psi_n(x) \propto e^{ikx}$

- d. $\psi_n(x) \propto x^n$
 - e. $\psi_n(x) \propto \delta(x - n)$
-

6. Which of the following are the energy eigenvalues E_n for the particle in a box?

- a. $E_n = \frac{\hbar^2 n^2}{2ma^2}$
 - b. $E_n = \frac{\hbar^2 n^2}{8ma^2}$
 - c. $E_n = \frac{1}{2}mn^2a^2$
 - d. $E_n = \frac{n\hbar c}{a}$
 - e. $E_n = n\hbar\omega$
-

7. The harmonic oscillator eigenfunctions are expressed in terms of which special functions?

- a. Legendre polynomials P_n
 - b. Laguerre polynomials L_n
 - c. Hermite polynomials H_n
 - d. Bessel functions J_n
 - e. Spherical harmonics Y_ℓ^m
-

8. Which of the following are the energy levels of the harmonic oscillator?

a. $E_n = \hbar\omega \left(n + \frac{1}{2}\right)$

b. $E_n = \hbar\omega n$

c. $E_n = \frac{\hbar^2 n^2}{8ma^2}$

d. $E_n = \frac{\hbar^2 n^2}{2m}$

e. $E_n = \frac{1}{2}kx^2$

9. What does the rigid rotor model represent?

- a. A particle moving in a 1D infinite well
 - b. Two point masses at fixed distance rotating in 3D
 - c. A mass on a spring moving in 1D
 - d. A free particle in 3D
 - e. A charged particle in a uniform electric field
-

10. Which of the following are the energy levels of the rigid rotor?

a. $E_\ell = \frac{\hbar^2}{2I} \ell(\ell + 1)$

b. $E_\ell = \hbar\omega(\ell + \frac{1}{2})$

c. $E_\ell = \frac{\hbar^2 \ell^2}{8I}$

d. $E_\ell = \frac{\hbar^2}{2I} m^2$

e. $E_\ell = \frac{1}{2}I\omega^2$

Answers: 1.c, 2.b, 3.e, 4.d, 5.a, 6.b, 7.c, 8.a, 9.b, 10.a

Chapter 5

The Hydrogen Atom

In this chapter we will consider the hydrogen atom as a proton fixed at the origin, orbited by an electron of reduced mass μ . The potential due to electrostatic attraction is:

$$V(r) = -\frac{e^2}{4\pi\varepsilon_0 r}, \quad (5.1)$$

where ε_0 is the constant permittivity of vacuum. The kinetic energy term in the Hamiltonian is

$$K = -\frac{\hbar^2}{2\mu} \nabla^2, \quad (5.2)$$

where ∇^2 is the Laplace operator (*Laplacian*) representing the divergence of the gradient of a function. Recall that in 1-dimension the kinetic energy is proportional to the second derivative of the wave function with respect to the position. In 3-dimension, the first derivative along all three dimension of space is called gradient, which is written in cartesian coordinates $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$. The Laplacian is the divergence $\nabla \cdot$ of the gradient (effectively, it replaces the second derivatives in the 1-D case), and can be written in cartesian coordinates as $\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. The TISEq for the Hydrogen atom is therefore:

$$\left(-\frac{\hbar^2}{2\mu} \nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r} \right) \psi(r, \theta, \phi) = E\psi(r, \theta, \phi), \quad (5.3)$$

which, replacing the Laplacian in spherical coordinates, becomes:

$$-\frac{\hbar^2}{2\mu} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} \right] - \frac{e^2}{4\pi\varepsilon_0 r} \psi = E\psi. \quad (5.4)$$

This equation seems very complicated, but comparing the term in between square brackets with the TISEq of the rigid rotor, we immediately see some connections. Eq. (5.4) is a separable, partial differential equation that can be solved by factorizing the wave function $\psi(r, \theta, \phi)$ into $R_{nl}(r)Y_\ell^{m_\ell}(\theta, \phi)$, where $Y_\ell^{m_\ell}(\theta, \phi)$ are again the spherical harmonics that solved the TISEq for the rigid rotor. The radial part $R(r)$ obeys the equation:

$$-\frac{\hbar^2}{2\mu r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) \left[\frac{\hbar^2 \ell(\ell+1)}{2\mu r^2} + V(r) - E \right] R(r) = 0, \quad (5.5)$$

which is called the radial equation for the hydrogen atom. The solutions of the radial part are:

$$R_{n\ell}(r) = - \left[\frac{(n-\ell-1)!}{2n[(n+\ell)!]^3} \right]^{1/2} \left(\frac{2}{na_0} \right)^{\ell+3/2} r^\ell e^{-r/na_0} L_{n+\ell}^{2\ell+1} \left(\frac{2r}{na_0} \right) \quad (5.6)$$

where $0 \leq \ell \leq n-1$, and $a_0 = \frac{\varepsilon_0 h^2}{\pi \mu e^2} \approx 0.529 \text{ \AA}$ is the Bohr radius. The functions $L_{n+\ell}^{2\ell+1} \left(\frac{2r}{na_0} \right)$ are the associated Laguerre polynomials

The hydrogen atom eigenfunctions are:

$$\begin{aligned} \psi_{nlm_\ell}(r, \theta, \phi) &= R_{n\ell}(r)Y_\ell^{m_\ell}(\theta, \phi) = \\ &= - \left[\frac{(n-\ell-1)!}{2n[(n+\ell)!]^3} \right]^{1/2} \left(\frac{2}{na_0} \right)^{\ell+3/2} r^\ell e^{-r/na_0} L_{n+\ell}^{2\ell+1} \left(\frac{2r}{na_0} \right) Y_\ell^{m_\ell}(\theta, \phi) \end{aligned} \quad (5.7)$$

The quantum numbers n, ℓ, m_ℓ can take the following values:

$n = 1, 2, 3, \dots, \infty$ (principal quantum number),

$\ell = 0, 1, 2, \dots, n-1$ (azimuthal quantum number),

$m_\ell = -\ell, \dots, \ell$ (magnetic quantum number).

These functions are called the **hydrogen atom orbitals**, and are usually first encountered in introductory chemistry textbooks. Notice that—by definition—an orbital is a complex function (i.e., it has both a real and an imaginary component) that describes *exclusively* one electron.

The hydrogen atom eigenvalues are:

$$E_n = -\frac{1}{n^2} \frac{e^2}{8\pi\varepsilon_0 a_0} \quad (n = 1, 2, \dots, \infty). \quad (5.8)$$

Notice how the eigenvalues (i.e., the energy spectrum) do not depend on the azimuthal and magnetic quantum numbers, ℓ and m_ℓ . Energy levels with the same n , but different ℓ and/or m_ℓ are degenerate. This is, unfortunately, source of some oversimplification in most general chemistry textbook:

- According to the solution of the TISEq, the $2s$ and $2p$ orbitals of the hydrogen atom have the same energy.

In practice, this is not exactly correct either, because of a tiny effect called the Lamb shift. The description of this effect requires to go beyond the Schrödinger equation into the field of quantum electrodynamics. The Lamb shift, however, is not what is usually depicted in general chemistry textbook as the $2s - 2p$ energy difference. This tiny effect, however, is not the responsible for the shift that is usually shown in most general chemistry textbooks. Such difference is usually discussed in the context of the aufbau principle and does not apply to hydrogen. In other words, the $2s - 2p$ split is purely a many-electron effect, as we will discuss in chapter 10. The energy levels of the hydrogen atom are plotted in figure 5.1.

5.1 Wave Function Analysis

5.1.1 The radial wave functions

The solutions to the radial equation, eq xxx, involve two key components: an exponential decay factor e^{-r/na_0} and a polynomial factor built from the associated Laguerre polynomials, $L_{n+\ell}^{2\ell+1}\left(\frac{2r}{na_0}\right)$. The complete radial wave functions for the first three energy levels are:

$$R_{10}(r) = 2\left(\frac{1}{a_0}\right)^{3/2} e^{-r/a_0}.$$

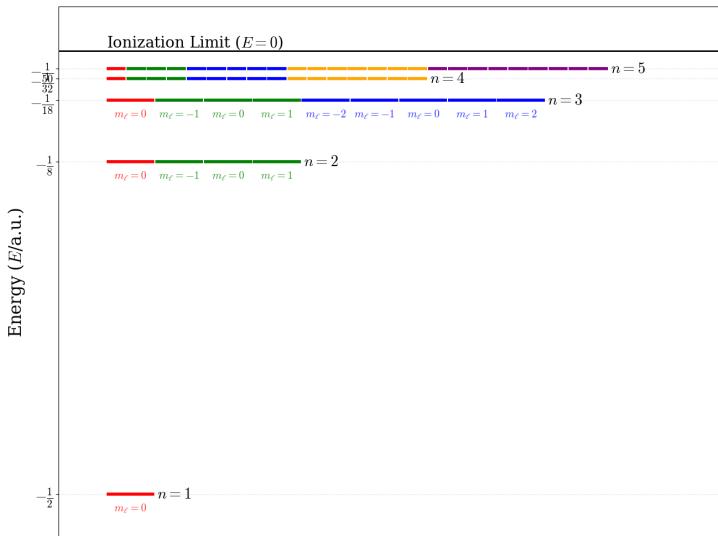


Figure 5.1: Energy level diagram and degeneracy pattern for the first five levels of the hydrogen atom.

$$\begin{aligned}
R_{20}(r) &= \frac{1}{2\sqrt{2}} \left(\frac{1}{a_0}\right)^{3/2} \left(2 - \frac{r}{a_0}\right) e^{-r/2a_0}. \\
R_{21}(r) &= \frac{1}{2\sqrt{6}} \left(\frac{1}{a_0}\right)^{3/2} \left(\frac{r}{a_0}\right) e^{-r/2a_0}. \\
R_{30}(r) &= \frac{2}{27\sqrt{3}} \left(\frac{1}{a_0}\right)^{3/2} \left[27 - 18\left(\frac{r}{a_0}\right) + 2\left(\frac{r}{a_0}\right)^2\right] e^{-r/3a_0}. \\
R_{31}(r) &= \frac{8}{27\sqrt{6}} \left(\frac{1}{a_0}\right)^{3/2} \frac{r}{a_0} \left(1 - \frac{r}{6a_0}\right) e^{-r/3a_0}. \\
R_{32}(r) &= \frac{1}{81\sqrt{30}} \left(\frac{1}{a_0}\right)^{3/2} \left(\frac{r}{a_0}\right)^2 e^{-r/3a_0}.
\end{aligned}$$

These formulas illustrate two critical features: (1) the exponential decay avoids the function to diverge as $r \rightarrow \infty$ confining the electron with higher probability closer to the nucleus, and (2) the polynomial factors give to each state a characteristic nodal structure. The radial wave functions are plotted in figure 5.2. Notice that R_{10} for the 1s state is always positive and decays smoothly—this orbital has no radial nodes. By contrast, R_{20} for the 2s state changes sign, crossing zero at $r = 2a_0$ (the radial node). This is a mathematical manifestation of increasing angular momentum confinement: higher energy states must oscillate more to fit their energy eigenvalues. The 2p radial wave function R_{21} starts at zero (due to the factor of r in the polynomial) and exhibits a single lobe before decaying. The total number of nodes in any orbital is $n - 1$. This count is split between radial and angular contributions:

- **Radial nodes:** $n - \ell - 1$ spherical surfaces where $R_{n\ell}(r) = 0$.
- **Angular nodes:** ℓ nodal planes or cones.

For example, the 2s orbital ($n = 2, \ell = 0$) has one total node, which is entirely radial (the sphere at $r = 2a_0$). The 2p orbital with $n = 2$ and $\ell = 1$ has one node, which is angular (the plane perpendicular to the orbital axis). The 3d orbital with $n = 3$ and $\ell = 2$ has $3 - 1 = 2$ total nodes, both angular (two nodal cones), with no radial nodes.

While the radial wave function $R_{n\ell}(r)$ describes the amplitude of the electron probability at distance r , what we typically measure experimentally is the radial probability distribution:

$$P(r) = 4\pi r^2 |R_{n\ell}(r)|^2. \quad (5.9)$$

This product of r^2 and $|R_{n\ell}|^2$ is crucial for physical interpretation and is plotted for the first six radial wave functions in figure 5.3. The factor $4\pi r^2$

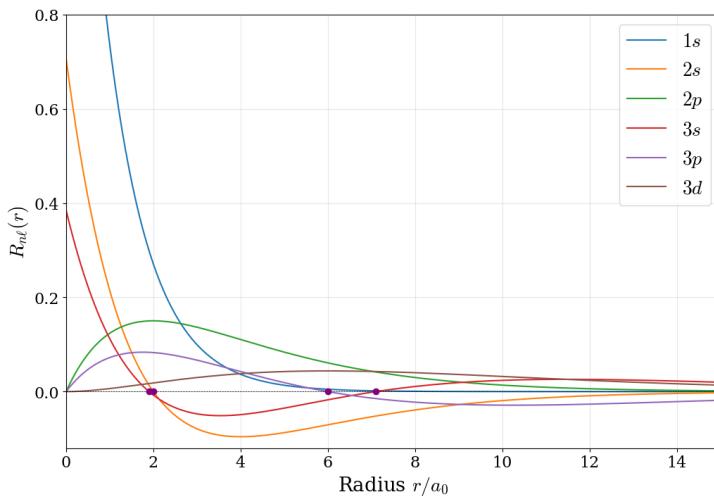


Figure 5.2: The radial wave functions of the hydrogen atom orbitals in the first three energy levels. The radial nodes are highlighted with purple dots.

represents the *surface area of a spherical shell* at distance r . Even though the probability density $|R_{nl}|^2$ is largest near the nucleus (often zero right at $r = 0$ for $\ell > 0$), the shell volume increases with r^2 , creating a maximum in $P(r)$ at an intermediate distance. This is why the most probable radius for the 1s electron is exactly the Bohr radius a_0 .

The radial probability plot reveals the structure of orbitals at different energy levels. For 1s, the single smooth peak centered at $a_0 \approx 0.53 \text{ \AA}$ shows why the ground state is compact and stable. For 2s, two peaks at $r/a_0 \approx 0.76 \text{ \AA}$ and 5.24 \AA are separated by a radial node at $r = 2a_0$. The presence of multiple peaks indicates that the electron, though more energetic (and thus less tightly bound), still has a non-zero probability of being found relatively close to the nucleus. The 3s orbital shows three peaks—the pattern continues with each increase in principal quantum number.

This radial probability distribution is the **observable quantity**—it represents the answer to the experimental question: “If I measure the distance of the electron from the nucleus, what are the odds of finding it between radius r and $r + dr$? ”

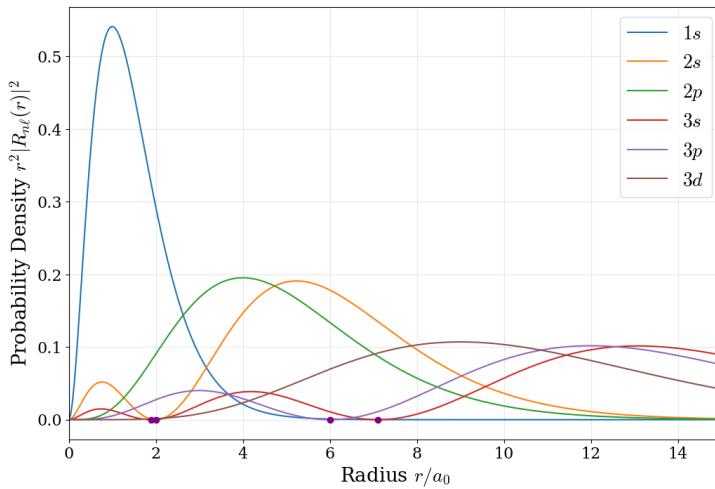


Figure 5.3: The radial probability functions of the hydrogen atom orbitals in the first three energy levels. The radial nodes are highlighted with purple dots.

5.1.2 The angular wave functions

As we already discussed, the angular part of the Schrödinger equation is solved by the same spherical harmonics we discussed for the solution of the particle on a sphere in Chapter 4. The complete complex spherical harmonics are eigenfunctions of the magnitude of the total angular momentum, \mathbf{L}^2 , and of one of its component along an arbitrary axis, L_z . These two conditions determine the **physical meaning** of the quantum numbers ℓ and m_ℓ . They also establish why the spherical harmonics create the characteristic orbital shapes:

- **n (principal quantum number):** Determines the energy $E_n = -13.6 \text{ eV}/n^2$ and the size of the orbital.
- **ℓ (orbital angular momentum quantum number):** Ranges from 0 to $n - 1$ and determines orbital shape (s for $\ell = 0$, p for $\ell = 1$, d for $\ell = 2$, f for $\ell = 3$, and so on.) The “shape of the orbital” is connected to how the probability density varies with the (angular) direction at a given radius. That information is entirely in the angular part of the wave function and is hence inherited from the shape of the spherical harmonics described in Chapter 4.
- **m_ℓ (magnetic quantum number):** Ranges from $-\ell$ to $+\ell$, specifies the orientation of the orbital in space. The quantum number ℓ quantizes the magnitude of the total angular momentum, but its direction is uncertain, except for one component along an arbitrary axis. By convention we take the z -axis and quantize its component via the quantum number m_ℓ . Therefore m_ℓ determines the orientation along the associated z -axis.

The energy depends *only* on n , making all orbitals with the same n but different ℓ values degenerate. This is a special feature of the Coulomb potential; most other potentials would split these energy levels.

Visualization of the complex spherical harmonics can happen in many different ways, as we have already shown in the particle-on-a-sphere section of Chapter 4. The difficulties associated with the visualization of the complex spherical harmonics translate directly to the visualization of the atomic orbitals, which we will discuss next.

5.1.3 Building the complete orbital

The complete wave function given by the multiplication of the radial components and the complex spherical harmonic encodes *all* the quantum infor-

mation about the electron's state. The radial part captures the probability of finding the electron at a given distance r (independent of direction), while the angular part specifies how this probability is distributed among different directions.

Since orbitals are functions of three spatial dimensions, visualizing their amplitude would require a four-dimensional plot. One way to reduce the represent them in 3D is to map the amplitude to a color map, similarly to figures 4.7 and 4.8 in Chapter 4. This, however, doesn't solve the issue with the visualization of complex functions, which would require two different 3D plots, one for the real part of the wave function (the amplitude) and one for the imaginary part (the phase). These sets of complementary plots are reported in figure 5.4 for the real part of the hydrogen atom orbitals and in figure 5.5 for the imaginary part.

In chemistry, we typically use *real orbitals* instead of the pure complex eigenstates. This is convenient because they align with molecular symmetry, bonding intuition, and computational visualization. The real orbitals are obtained from linear combinations of spherical harmonics with opposite m_ℓ values,¹ for example:

$$\begin{aligned}\psi_{2p_z} &= \psi_{2,1,0} \quad (\text{already real}), \\ \psi_{2p_x} &= \frac{1}{\sqrt{2}} (\psi_{2,1,+1} + \psi_{2,1,-1}), \\ \psi_{2p_y} &= \frac{i}{\sqrt{2}} (\psi_{2,1,+1} - \psi_{2,1,-1}).\end{aligned}\tag{5.10}$$

These combinations eliminate the azimuthal phase dependence, producing three real-valued orbitals with the familiar dumbbell shapes oriented along the x , y , and z axes. Each real orbital still represents a valid quantum state—they are simply a different choice of *basis* for describing the same subspace of quantum states. The real p orbitals are superpositions of definite angular momentum states, so they don't have definite m_ℓ values, but they have all other properties encoded in the original states.

The crucial point is: **the real and complex orbital representations are mathematically equivalent descriptions of the same physical system.** They provide different perspectives (real vs. complex) but encode identical measurable information.

¹The combination of complex spherical harmonics into real ones results in cubic (or tesseral) harmonics.

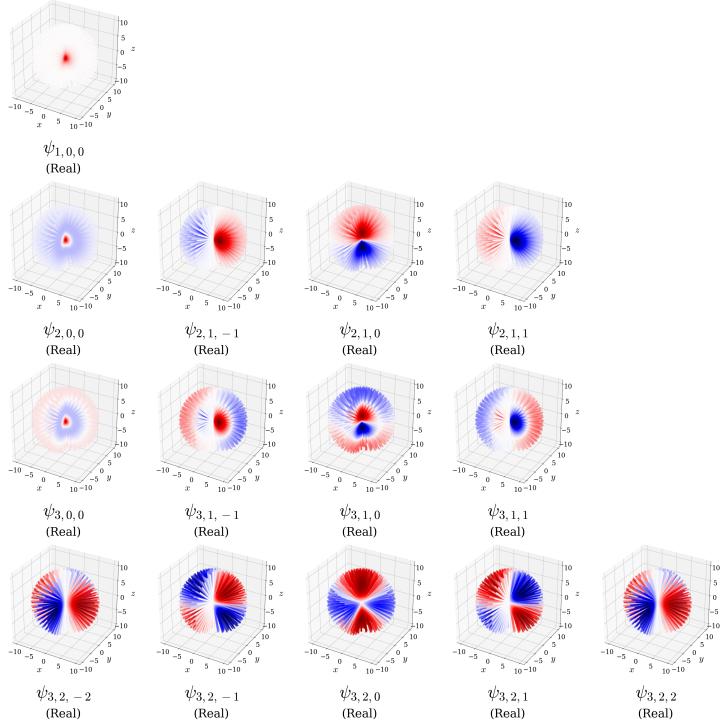


Figure 5.4: Color gradient maps of the real part ($\Re[\psi_{n\ell m_\ell}(r, \theta, \phi)]$) for hydrogen atom orbitals. Color scheme: Red (positive), blue (negative), white/light (nodal/grey) (where $\psi_{n\ell m_\ell} = 0$).

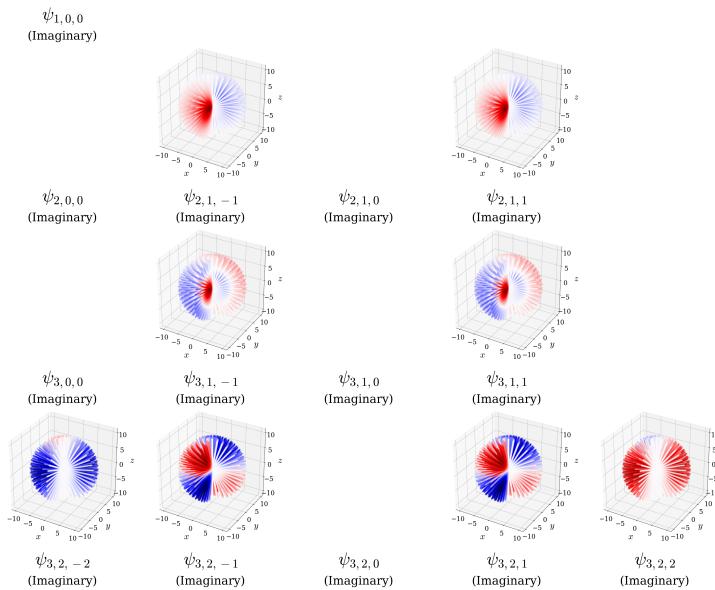


Figure 5.5: Color gradient maps of the imaginary part ($\Im[\psi_{n\ell m_\ell}(r, \theta, \phi)]$) for hydrogen atom orbitals. Color scheme: Red (positive), blue (negative), white/light (nodal grat circles where $\psi_{n\ell m_\ell} = 0$).

One final comment on the nature of orbital representations merits particular emphasis here. Pictures of orbitals such as those shown in figures 5.4 and 5.5 carry **no direct physical significance**. These pictures simply plot mathematical functions. While they can help infer electron probability distributions, they should never be interpreted as literal “regions of space” where electrons reside. Unfortunately, renderings of these figures are commonly misrepresented as actual orbital images, and students are often led to believe that these pictures represent an actual, measurable, physical reality. **They do not!** Since the probability density is the quantity that can be experimentally measured, $|\psi_{n\ell m_\ell}|^2$ is the only representation that has physical significance. These “electron density probability clouds” can be visualized in several ways:

1. **Probability density contour surface plots (isodensity maps):** Show an isosurface enclosing (e.g.) 80% of the electron probability. This is done for a few representative orbitals in figure 5.6. These plots are excellent for conveying orbital shape but loses information about the probability gradient.
2. **Volumetric density plots:** Use opacity to represent $|\psi|^2$ throughout a region. This is done for a few representative orbitals in figure 5.7. These plots preserve all spatial information but can be harder to interpret visually.
4. **Dot-density plots:** Scatter dots with density proportional to $|\psi|^2$. This is done for a few representative orbitals in figure 5.8. These plots represent a stochastic “electron cloud” model where the actual spatial density of the dots represents the probability of finding the electron and are very intuitive for understanding probability as a statistical distribution.

Each method highlights different aspects of the orbital structure. The key is understanding that all these representations show the *same* information (the probability density) through different visual encodings. Finally, it is important to note that most chemistry textbooks depict orbital symmetry on probability density diagrams, such as those in Figures 5.6, 5.7, and 5.8, by overlaying the signs of the wave functions using colors or shading. This common practice is misleading, as these plots represent the probability of finding the electron, which must be non-negative everywhere. Nevertheless,

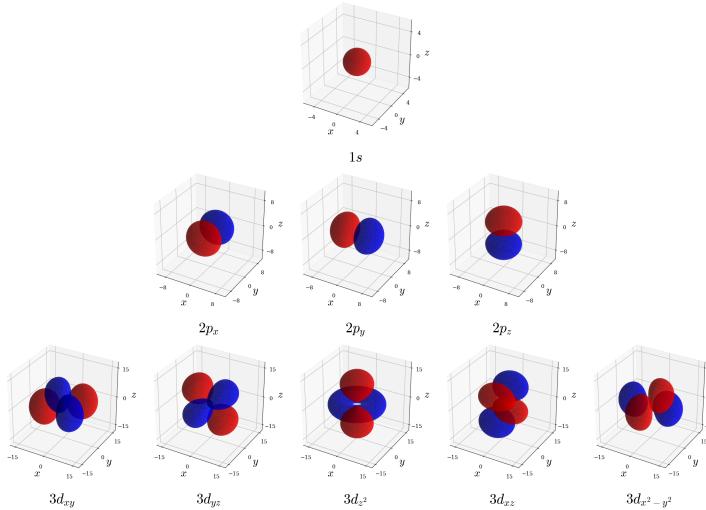


Figure 5.6: Probability density contour plots (isosurfaces) showing the isosurface containing 80% of the electron probability ($|\psi_{n\ell m_\ell}(r, \theta, \phi)|^2$) for a few representative hydrogen atom orbitals. Regions are colored blue (positive wave function phase) and red (negative phase) following standard chemistry convention; note this shading illustrates $\psi_{n\ell m_\ell}$ sign only, not probability itself which remains non-negative everywhere.

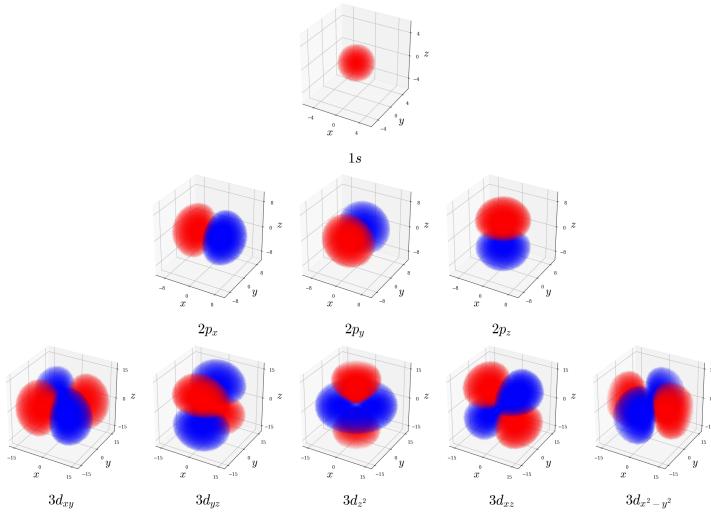


Figure 5.7: Volumetric cloud plots for a few representative hydrogen atom orbitals. Regions are colored blue (positive wave function phase) and red (negative phase) following standard chemistry convention; note this shading illustrates $\psi_{n\ell m_\ell}$ sign only, not probability itself which remains non-negative everywhere.

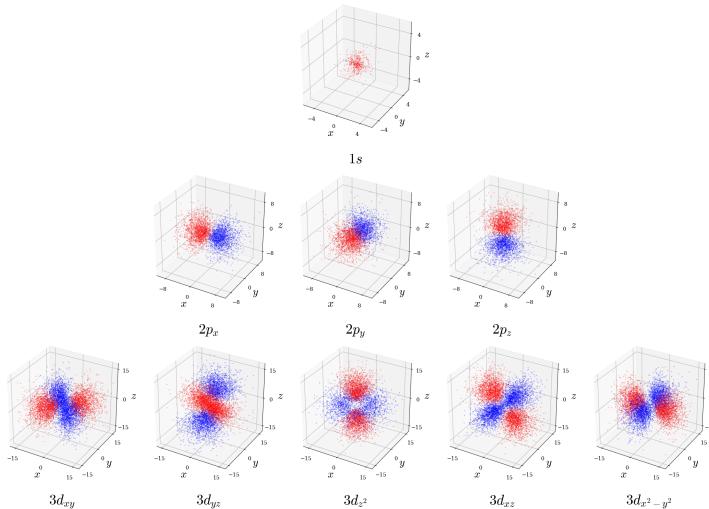


Figure 5.8: Dot-density plots for a few representative hydrogen atom orbitals. Individual dots are colored blue (positive wave function phase) and red (negative phase) following standard chemistry convention; note this shading illustrates $\psi_{n\ell m_\ell}$ sign only, not probability itself which remains non-negative everywhere.

because this convention is too pervasive in the literature to ignore, we retain it here to familiarize students with standard representations while urging them to recognize its limitations.

5.1.4 Experimental Access to Orbitals

A common question is: “Can we actually measure an orbital?” The answer is nuanced. We cannot directly observe $\psi_{n\ell m_\ell}(\mathbf{r})$ because the wave function is not an observable—quantum mechanics predicts only probabilities for measurement outcomes. However, we can:

- Measure the *probability density* through many repeated experiments (electron microscopy, photoelectron spectroscopy).
- Recover both amplitude and phase information through sophisticated techniques like Angle-Resolved Photoemission Spectroscopy (ARPES) combined with iterative reconstruction algorithms.
- Verify orbital shapes indirectly through spectroscopic measurements, magnetic susceptibility, x-ray diffraction, and other bulk properties.

Recent attosecond and femtosecond experiments have directly imaged the amplitude of molecular orbitals and even tracked the phase evolution during chemical reactions, confirming that quantum mechanical orbital descriptions accurately reflect physical reality.

5.2 Chapter Review

5.2.1 Study Questions

1. Which of the following is the electrostatic potential energy, $V(r)$, in the hydrogen atom?

- $V(r) = -\frac{e^2}{4\pi\epsilon_0 r}$
- $V(r) = +\frac{e^2}{4\pi\epsilon_0 r}$
- $V(r) = \frac{1}{2}kr^2$

d. $V(r) = -G \frac{m_p m_e}{r}$

e. $V(r) = 0$

2. Which of the following is the kinetic-energy operator used in the hydrogen Hamiltonian?

a. $K = -\frac{\hbar^2}{2\mu} \nabla^2$

b. $K = -\frac{\hbar^2}{2m_e} \frac{d^2}{dr^2}$

c. $K = \frac{p^2}{2\mu}$

d. $K = -\frac{\hbar^2}{2} \nabla$

e. $K = -\frac{\hbar^2}{2M} \nabla^2$

3. How do you separate the hydrogen-atom wave function in spherical coordinates?

a. $\psi(r, \theta, \phi) = f(r)g(\theta)$

b. $\psi(r, \theta, \phi) = R_{n\ell}(r) Y_\ell^{m_\ell}(\theta, \phi)$

c. $\psi(r, \theta, \phi) = R(r) \cos \theta$

d. $\psi(r, \theta, \phi) = Y_\ell^{m_\ell}(\theta, \phi)$

e. $\psi(r, \theta, \phi) = R_{n\ell}(r) e^{ikr}$

4. Which special functions give the angular part of the hydrogen-atom wave functions?

- a. Hermite polynomials
 - b. Associated Laguerre polynomials
 - c. Bessel functions
 - d. Plane waves
 - e. Spherical harmonics
-

5. How do you calculate the Bohr's radius?

- a. $a_0 = \frac{\varepsilon_0 h^2}{\pi \mu e^2}$
 - b. $a_0 = \frac{4\pi \varepsilon_0 \hbar^2}{\mu e^2}$
 - c. $a_0 = \frac{\hbar^2}{\mu e^2}$
 - d. $a_0 = \frac{\varepsilon_0 h}{\mu e}$
 - e. $a_0 = \frac{e^2}{4\pi \varepsilon_0 \hbar^2}$
-

6. Which special functions give the radial part of the hydrogen-atom wave functions?

- a. Hermite polynomials
- b. Legendre polynomials
- c. Associated Laguerre functions
- d. Spherical harmonics
- e. Bessel functions

7. What are the allowed values of the principal quantum number n ?

- a. $n = 0, 1, 2, \dots$
 - b. $n = -\infty \dots \infty$
 - c. $n = 1, 2, 3, \dots$
 - d. $n = \frac{1}{2}, \frac{3}{2}, \dots$
 - e. $n = 2, 4, 6, \dots$
-

8. For a given n , what are the allowed values of the azimuthal quantum number ℓ ?

- a. $\ell = 0, 1, 2, \dots, n - 1$
 - b. $\ell = 1, 2, \dots, n$
 - c. $\ell = -n, \dots, n$
 - d. $\ell = 0, \pm 1, \pm 2, \dots$
 - e. $\ell = n$ only
-

9. For or a given ℓ , what are the allowed values of the magnetic quantum number m_ℓ ?

- a. $m_\ell = 0, 1, \dots, \ell$
- b. $m_\ell = 1, 2, \dots, \ell$

- c. $m_\ell = -n, \dots, n$
 - d. $m_\ell = \pm \frac{1}{2}$
 - e. $m_\ell = -\ell, -\ell + 1, \dots, \ell$
-

10. Which statement about orbitals does the chapter explicitly emphasize?

- a. Orbitals physically contain up to two electrons
 - b. Each orbital is the mathematical description of one and only one electron
 - c. Orbitals are purely classical trajectories
 - d. Orbitals only exist for many-electron atoms
 - e. Orbitals always have dumbbell shapes in real space
-

Answers: 1.a, 2.a, 3.b, 4.e, 5.a, 6.c, 7.c, 8.a, 9.e, 10.b

Chapter 6

Operators and Mathematical Background

So far, we have seen a few simple examples of how to solve the TISEq. For the general case, the mathematical formulation of quantum mechanics is built upon the concept of an operator. An operator is a function over a space of physical states onto another space of physical states. Operators do not exist exclusively in quantum mechanics, but they can also be used in classical mechanics. In chapter 2, we have seen at least a couple of them, namely the Lagrangian, L , and Hamiltonian, H . In quantum mechanics, however, the concept of an operator is the basis of the complex mathematical treatment that is necessary for more complicated cases. In this chapter, we will discuss the mathematics of quantum mechanical operators, and we will recast the results for the analytical cases in light of the new framework. As we will see, this framework is even simpler than what we have seen in the previous chapter. This simplicity, however, will open the door to the “stranger” side of quantum mechanics.

6.1 Operators in Quantum Mechanics

The central concept in this new framework of quantum mechanics is that every observable (i.e., any quantity that can be measured in a physical experiment) is associated with an operator. To distinguish between classical mechanics operators and quantum mechanical ones, we use a hat symbol $\hat{}$

on top of the latter. Physical pure states in quantum mechanics are represented as unit-norm (probabilities are normalized to one) vectors in a special complex Hilbert space. Following the definition, an operator is a function that projects a vector in the Hilbert space onto the space of physical observables. Since observables are values that come up as the result of the experiment, quantum mechanical operators must yield real eigenvalues.¹ Operators that possess this property are called *Hermitian*. In the wave mechanics formulation of quantum mechanics that we have seen so far, the wave function varies with space and time—or equivalently momentum and time—and observables are differential operators. A completely analogous formulation is possible in terms of matrices. In the matrix formulation of quantum mechanics, the norm of the physical state should stay fixed, so the evolution operator should be unitary, and the operators can be represented as matrices.

The expectation value of an operator \hat{A} for a system with wave function $\psi(\mathbf{r})$ living in a Hilbert space with unit vector \mathbf{r} (i.e., in three-dimensional Cartesian space $\mathbf{r} = \{x, y, z\}$), is given by:

$$\langle A \rangle = \int \psi^*(\mathbf{r}) \hat{A} \psi(\mathbf{r}) d\mathbf{r}, \quad (6.1)$$

and if \hat{A} is a Hermitian operator, all physical observables are represented by such expectation values. It is easy to show that if \hat{A} is a linear operator with an eigenfunction g , then any multiple of g is also an eigenfunction of \hat{A} .

6.1.1 Basic Properties of Operators

Most of the properties of operators are obvious, but they are summarized below for completeness. The sum and difference of two operators \hat{A} and \hat{B} are given by:

$$\begin{aligned} (\hat{A} + \hat{B})f &= \hat{A}f + \hat{B}f \\ (\hat{A} - \hat{B})f &= \hat{A}f - \hat{B}f. \end{aligned} \quad (6.2)$$

The product of two operators is defined by:

¹But they might not be strictly real.

$$\hat{A}\hat{B}f \equiv \hat{A}[\hat{B}f] \quad (6.3)$$

Two operators are equal if

$$\hat{A}f = \hat{B}f \quad (6.4)$$

for all functions f . The identity operator $\hat{1}$ does nothing (or multiplies by 1):

$$\hat{1}f = f \quad (6.5)$$

The associative law holds for operators:

$$\hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C} \quad (6.6)$$

The commutative law does not generally hold for operators. In general, $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. It is convenient to define the quantity:

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \quad (6.7)$$

which is called the **commutator** of \hat{A} and \hat{B} . Note that the order matters, so that $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$. If \hat{A} and \hat{B} happen to commute, then $[\hat{A}, \hat{B}] = 0$.

6.1.2 Linear Operators

Almost all operators encountered in quantum mechanics are linear. A linear operator is any operator \hat{A} satisfying the following two conditions:

$$\begin{aligned} \hat{A}(f + g) &= \hat{A}f + \hat{A}g, \\ \hat{A}(cf) &= c\hat{A}f, \end{aligned} \quad (6.8)$$

where c is a constant and f and g are functions. As an example, consider the operators $\frac{d}{dx}$ and $(\)^2$. We can see that $\frac{d}{dx}$ is a linear operator because:

$$\begin{aligned} \frac{d}{dx}[f(x) + g(x)] &= \frac{d}{dx}f(x) + \frac{d}{dx}g(x), \\ \frac{d}{dx}[cf(x)] &= c(d/dx)f(x). \end{aligned} \quad (6.9)$$

However, $(\cdot)^2$ is not a linear operator because:

$$(f(x) + g(x))^2 \neq (f(x))^2 + (g(x))^2 \quad (6.10)$$

6.1.3 Hermitian Operators

Hermitian operators are characterized by the self-adjoint property:

$$\int \psi_a^* (\hat{A}\psi_a) d\mathbf{r} = \int \psi_a (\hat{A}\psi_a)^* d\mathbf{r}, \quad (6.11)$$

where the integral is performed over all space. This property guarantees that all the eigenvalues of the operators are real. Defining a as the eigenvalue of operator \hat{A} using:

$$\hat{A}\psi(\mathbf{r}) = a\psi(\mathbf{r}), \quad (6.12)$$

we can prove that a is real by replacing eq. (6.12) into eq. (6.11):

$$\begin{aligned} a \int \psi_a^* \psi_a d\mathbf{r} &= a^* \int \psi_a \psi_a^* d\mathbf{r} \\ (a - a^*) \int |\psi_a|^2 d\mathbf{r} &= 0, \end{aligned} \quad (6.13)$$

and since $|\psi_a|^2$ is never negative, either $a = a^*$ or $\psi_a = 0$. Since $\psi_a = 0$ is not an acceptable wavefunction, $a = a^*$, and a is real.

The following additional properties of Hermitian operators can also be proven with some work:

$$\int \psi^* \hat{A}\psi d\mathbf{r} = \int (\hat{A}\psi)^* \psi d\mathbf{r}, \quad (6.14)$$

and for any two states ψ_1 and ψ_2 :

$$\int \psi_1^* \hat{A}\psi_2 d\mathbf{r} = \int (\hat{A}\psi_1)^* \psi_2 d\mathbf{r}. \quad (6.15)$$

Taking ψ_a and ψ_b as eigenfunctions of \hat{A} with eigenvalues a and b with $a \neq b$, and using eq. (6.15), we obtain:

$$\begin{aligned} \int \psi_a^* \hat{A} \psi_b d\mathbf{r} &= \int (\hat{A} \psi_a)^* \psi_b d\mathbf{r} \\ b \int \psi_a^* \psi_b d\mathbf{r} &= a^* \int \psi_a^* \psi_b d\mathbf{r} \\ (b - a) \int \psi_a^* \psi_b d\mathbf{r} &= 0. \end{aligned} \quad (6.16)$$

Thus, since $a = a^*$, and since we assumed $b \neq a$, we must have $\int \psi_a^* \psi_b d\mathbf{r} = 0$, i.e. ψ_a and ψ_b are orthogonal. In other words, eigenfunctions of a Hermitian operator with different eigenvalues are orthogonal (or can be chosen to be so).

6.2 Eigenfunctions and Eigenvalues

As we have already seen, an eigenfunction of an operator \hat{A} is a function f such that the application of \hat{A} on f gives f again, times a constant:

$$\hat{A}f = kf, \quad (6.17)$$

where k is a constant called the eigenvalue.

When a system is in an eigenstate of observable A (i.e., when the wave function is an eigenfunction of the operator \hat{A}) then the expectation value of A is the eigenvalue of the wave function. Therefore:

$$\hat{A}\psi(\mathbf{r}) = a\psi(\mathbf{r}), \quad (6.18)$$

then:

$$\begin{aligned} \langle A \rangle &= \int \psi^*(\mathbf{r}) \hat{A} \psi(\mathbf{r}) d\mathbf{r} \\ &= \int \psi^*(\mathbf{r}) a \psi(\mathbf{r}) d\mathbf{r} \\ &= a \int \psi^*(\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r} = a, \end{aligned} \quad (6.19)$$

which implies that:

$$\int \psi^*(\mathbf{r})\psi(\mathbf{r})d\mathbf{r} = 1. \quad (6.20)$$

This property of wave functions is called normalization, and in the one-electron TISEq guarantees that the maximum probability of finding an electron over the entire space is one.²

A unique property of quantum mechanics is that a wave function can be expressed not just as a simple eigenfunction, but also as a combination of several of them. We have in part already encountered such property in the previous chapter, where complex hydrogen orbitals have been combined to form corresponding linear ones. As a general example, let us consider a wave function written as a linear combination of two eigenstates of \hat{A} , with eigenvalues a and b :

$$\psi = c_a\psi_a + c_b\psi_b, \quad (6.21)$$

where $\hat{A}\psi_a = a\psi_a$ and $\hat{A}\psi_b = b\psi_b$. Then, since ψ_a and ψ_b are orthogonal and normalized (usually abbreviated as orthonormal), the expectation value of A is:

$$\begin{aligned} < A > &= \int \psi^* \hat{A} \psi d\mathbf{r} \\ &= \int [c_a\psi_a + c_b\psi_b]^* \hat{A} [c_a\psi_a + c_b\psi_b] d\mathbf{r} \\ &= \int [c_a\psi_a + c_b\psi_b]^* [ac_a\psi_a + bc_b\psi_b] d\mathbf{r} \\ &= a|c_a|^2 \int \psi_a^* \psi_a d\mathbf{r} + bc_a^* c_b \int \psi_a^* \psi_b d\mathbf{r} + ac_b^* c_a \int \psi_b^* \psi_a d\mathbf{r} + b|c_b|^2 \int \psi_b^* \psi_b d\mathbf{r} \\ &= a|c_a|^2 + b|c_b|^2. \end{aligned} \quad (6.22)$$

This result shows that the average value of A is a weighted average of eigenvalues, with the weights being the squares of the coefficients of the eigenvectors in the overall wavefunction.³

²Imposing the normalization condition is the best way to find the constant A in the solution of the TISEq for the particle in a box, a topic that we delayed in chapter 4.

³This section was adapted in part from Prof. C. David Sherrill's A Brief Review of Elementary Quantum Chemistry Notes available here.

6.3 Common Operators in Quantum Mechanics

Some common operators occurring in quantum mechanics are collected in the table below:

Observable Name	Symbol	Operator	Operation
Position	\mathbf{r}	$\hat{\mathbf{r}}$	Multiply by \mathbf{r}
Momentum	\mathbf{p}	$\hat{\mathbf{p}}$	$-i\hbar \left(\hat{i} \frac{\partial}{\partial x} + \hat{j} \frac{\partial}{\partial y} + \hat{k} \frac{\partial}{\partial z} \right)$
Kinetic energy	K	\hat{K}	$-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right)$
Potential energy	$V(\mathbf{r})$	$\hat{V}(\mathbf{r})$	Multiply by $V(\mathbf{r})$
Total energy	E	\hat{H}	$-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V(\mathbf{r})$
Angular momentum	L	\hat{L}^2	$\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$
	L_x	\hat{L}_x	$-i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right)$
	L_y	\hat{L}_y	$-i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right)$
	L_z	\hat{L}_z	$-i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)$

In the sections below we analyze in details two main operators for the energy and the angular momentum.

6.3.1 Hamiltonian Operator

The main quantity that quantum mechanics is interested in is the total energy of the system, E . The operator corresponding to this quantity is called *Hamiltonian*:

$$\hat{H} = -\frac{\hbar^2}{2} \sum_i \frac{1}{m_i} \nabla_i^2 + V, \quad (6.23)$$

where i is an index over all the particles of the system. Using the formalism of operators in conjunction with eq. (6.23), we can write the TISEq just simply as:

$$\hat{H}\psi = E\psi. \quad (6.24)$$

Comparing eq. (6.23) to the classical analog in eq. (2.16), we notice how the first term in the Hamiltonian operator represents the corresponding kinetic

energy operator, \hat{K} , while the second term represents the potential energy operator, \hat{V} . For a one-electron system—such as the ones we studied in chapter 4—we can write:

$$\hat{K} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) = \frac{\hbar^2}{2m} \nabla^2, \quad (6.25)$$

which is universal and applies to all systems. The potential energy operator \hat{V} is what differentiates each system. Using eq. (6.24), we can then simply obtain the TISEq for each of the first three models discussed in chapter 4 by simply using:

$$\begin{aligned} \text{Free particle: } & \hat{V} = 0, \\ \text{Particle in a box: } & \hat{V} = 0 \text{ inside the box, } \hat{V} = \infty \text{ outside the box,} \\ \text{Harmonic oscillator: } & \hat{V} = \frac{1}{2} kx^2. \end{aligned} \quad (6.26)$$

While these three cases are trivial to solve, the case of the rigid rotor is more complicated to solve, since the kinetic energy operator needs to be solved in spherical polar coordinates, as we will show in the next section.

6.3.2 Angular Momentum Operator

To write the kinetic energy operator \hat{K} for the rigid rotor, we need to express the Laplacian, ∇^2 , in spherical polar coordinates:

$$\nabla^2 = \nabla_r^2 - \frac{\hat{L}^2}{r^2}, \quad (6.27)$$

where $\nabla_r^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right)$ is the radial Laplacian, and \hat{L}^2 is the square of the total angular momentum operator, which is:

$$\begin{aligned} \hat{L}^2 &= \hat{L} \cdot \hat{L} = (\mathbf{i}\hat{L}_x + \mathbf{j}\hat{L}_y + \mathbf{k}\hat{L}_z) \cdot (\mathbf{i}\hat{L}_x + \mathbf{j}\hat{L}_y + \mathbf{k}\hat{L}_z) \\ &= \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \end{aligned} \quad (6.28)$$

with $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ the unitary vectors in three-dimensional space. The component along each direction, $\{\hat{L}_x, \hat{L}_y, \hat{L}_z\}$, are then expressed in cartesian coordinates using the following formulas:

$$\begin{aligned}\hat{L}_x &= -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right), \\ \hat{L}_y &= -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right), \\ \hat{L}_z &= -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right).\end{aligned}\quad (6.29)$$

The eigenvalues equation corresponding to the total angular momentum is:

$$\hat{L}^2 Y(\theta, \varphi) = \hbar^2 \ell(\ell + 1) Y_\ell^{m_\ell}(\theta, \varphi), \quad (6.30)$$

where ℓ is the azimuthal quantum number and $Y_\ell^m(\theta, \varphi)$ are the spherical harmonics, both of which we already encountered in chapter 4. Recall once again that each energy level E_ℓ is $(2\ell + 1)$ -fold degenerate in m_ℓ , since m_ℓ can have values $-\ell, -\ell + 1, \dots, \ell - 1, \ell$. This means that there are $(2\ell + 1)$ states with the same energy E_ℓ , each characterized by the magnetic magnetic quantum number m_ℓ . This quantum number can be determined using the following eigenvalues equation:

$$\hat{L}_z Y(\theta, \varphi) = \hbar m_\ell Y_\ell^{m_\ell}(\theta, \varphi). \quad (6.31)$$

The interpretation of these results is rather complicated, since the angular momenta are quantum operators and they cannot be drawn as vectors like in classical mechanics. Nevertheless, it is common to depict them heuristically as in figure 6.1⁴, where a set of states with quantum numbers $\ell = 2$, and $m_\ell = -2, -1, 0, 1, 2$ are reported. Since $|L| = \sqrt{L^2} = \hbar\sqrt{6}$, the vectors are all shown with length $\hbar\sqrt{6}$. The rings represent the fact that L_z is known with certainty, but L_x and L_y are unknown; therefore every classical vector with the appropriate length and z -component is drawn, forming a cone. The expected value of the angular momentum for a given ensemble of systems in the quantum state characterized by ℓ and m_ℓ , could be somewhere on this cone but it cannot be defined for a single system.

⁴This diagram is taken from Wikipedia by user Maschen, and is of public domain

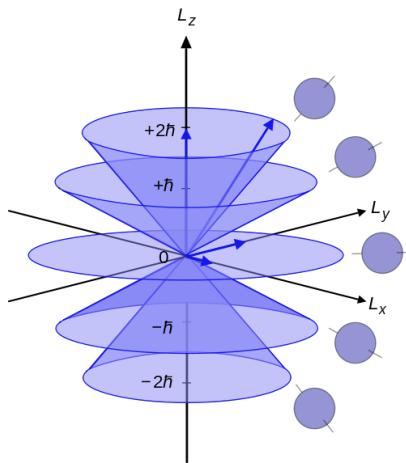


Figure 6.1: Illustration of the vector model of orbital angular momentum.

6.4 Chapter Review

6.4.1 Study Questions

1. Which expression gives the expectation value of an observable represented by operator \hat{a} in state $\psi(\mathbf{r})$?

- a. $\langle a \rangle = \int \psi^*(\mathbf{r}) \hat{a} \psi(\mathbf{r}) d\mathbf{r}$
- b. $\langle a \rangle = \int \hat{a} \psi^*(\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r}$
- c. $\langle a \rangle = \int |\psi(\mathbf{r})|^2 d\mathbf{r}$
- d. $\langle a \rangle = \int a \psi^*(\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r}$
- e. $\langle a \rangle = \int \psi(\mathbf{r}) \hat{a} \psi^*(\mathbf{r}) d\mathbf{r}$

2. An eigenfunction f of an operator \hat{a} with eigenvalue k satisfies which equation?

- a. $\hat{a}f = k$
 - b. $\hat{a}f = kf$
 - c. $\hat{a}f = f + k$
 - d. $\hat{a}^2f = kf$
 - e. $\hat{a}f = k^2f$
-

3. If f is an eigenfunction of \hat{a} with eigenvalue k , which statement is true for any nonzero scalar c ?

- a. cf has eigenvalue k/c
 - b. cf is an eigenfunction with eigenvalue ck
 - c. cf is no longer an eigenfunction
 - d. cf is an eigenfunction only if $|c| = 1$
 - e. cf is also an eigenfunction with the same eigenvalue k
-

4. What is the commutator of two operators \hat{a} and \hat{b} ?

- a. $[\hat{a}, \hat{b}] = \hat{a} + \hat{b}$
 - b. $[\hat{a}, \hat{b}] = \hat{a}\hat{b} + \hat{b}\hat{a}$
 - c. $[\hat{a}, \hat{b}] = \hat{a}\hat{b} - \hat{b}\hat{a}$
 - d. $[\hat{a}, \hat{b}] = \hat{a} - \hat{b}$
 - e. $[\hat{a}, \hat{b}] = 0$ by definition
-

5. Which important consequence follows for the eigenvalues of a hermitian operator \hat{a} ?

- a. All eigenvalues are real
 - b. All eigenvalues are purely imaginary
 - c. Eigenvalues occur only in complex-conjugate pairs
 - d. Eigenvalues must be integers
 - e. Eigenvalues must be positive
-

6. For a Hermitian operator, eigenfunctions corresponding to different eigenvalues have what property?

- a. They are always identical up to a phase
 - b. They are orthogonal (or can be chosen orthogonal)
 - c. They are linearly dependent
 - d. They have the same expectation value of all observables
 - e. They must be degenerate
-

7. In one dimension, the standard momentum operator \hat{p}_x acting on $\psi(x)$ is:

- a. $\hat{p}_x = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2}$
- b. $\hat{p}_x = i\hbar \frac{d}{dx}$
- c. $\hat{p}_x = -i\hbar \frac{d}{dx}$
- d. $\hat{p}_x = x$

e. $\hat{p}_x = \frac{d}{dx}$

8. If a wavefunction ψ is written as a linear combination of eigenfunctions $\{\psi_i\}$ of \hat{a} , $\psi = \sum_i c_i \psi_i$, what are the possible outcomes of a measurement of a ?

- a. Any complex number
 - b. Any real number in the spectrum of \hat{h}
 - c. Only the expectation value $\langle a \rangle$
 - d. Only the eigenvalues a_i corresponding to the eigenfunctions present in the expansion
 - e. Only the largest eigenvalue
-

9. In the same expansion $\psi = \sum_i c_i \psi_i$, what is the probability of obtaining eigenvalue a_k upon measuring a (assuming normalized eigenfunctions)?

- a. $|c_k|^2$
 - b. c_k
 - c. $|a_k|^2$
 - d. $1/|c_k|^2$
 - e. $\text{Re}(c_k)$
-

10. What is the general role of the Hamiltonian operator \hat{H} in quantum mechanics?

- a. It has no observable meaning

- b. It represents only the kinetic energy
 - c. It represents only the potential energy
 - d. It measures the position of particles
 - e. It represents the total energy and generates time evolution
-

Answers: 1.a, 2.b, 3.e, 4.c, 5.a, 6.b, 7.c, 8.d, 9.a, 10.e

Chapter 7

Spin

Spin is a special property of particles that has no classical analogue. Spin is an intrinsic form of angular momentum carried by elementary particles, such as the electron.

7.1 Stern-Gerlach Experiment

In 1920, Otto Stern and Walter Gerlach designed an experiment that unintentionally led to the discovery that electrons have their own individual, continuous spin even as they move along their orbital of an atom. The experiment was done by putting a silver foil in an oven to vaporize its atoms. The silver atoms were collected into a beam that passed through an inhomogeneous magnetic field. The result was that the magnetic beam split the beam into two (and only two) separate ones. The Stern–Gerlach experiment demonstrated that the spatial orientation of angular momentum is quantized into two components (up and down). Thus an atomic-scale system was shown to have intrinsically quantum properties. The experiment is normally conducted using electrically neutral particles such as silver atoms. This avoids the large deflection in the path of a charged particle moving through a magnetic field and allows spin-dependent effects to dominate.

If the particle is treated as a classical spinning magnetic dipole, it will precess in a magnetic field because of the torque that the magnetic field exerts on the dipole. If it moves through a homogeneous magnetic field, the forces exerted on opposite ends of the dipole cancel each other out and the trajectory of

the particle is unaffected. However, if the magnetic field is inhomogeneous then the force on one end of the dipole will be slightly greater than the opposing force on the other end, so that there is a net force which deflects the particle's trajectory. If the particles were classical spinning objects, one would expect the distribution of their spin angular momentum vectors to be random and continuous. Each particle would be deflected by an amount proportional to its magnetic moment, producing some density distribution on the detector screen. Instead, the particles passing through the Stern–Gerlach apparatus are equally distributed among two possible values, with half of them ending up at an upper spot (“spin up”), and the other half at the lower spot (“spin down”). Since the particles are deflected by a magnetic field, spin is a magnetic property that is associated to some intrinsic form of angular momentum. As we saw in chapter 6, the quantization of the angular momentum gives energy levels that are $(2\ell+1)$ -fold degenerate. Since along the direction of the magnet we observe only two possible eigenvalues for the spin, we conclude the following value for s :

$$2s + 1 = 2 \quad \Rightarrow \quad s = \frac{1}{2}. \quad (7.1)$$

The Stern-Gerlach experiment proves that electrons are spin- $\frac{1}{2}$ particles. These have only two possible spin angular momentum values measured along any axis, $+\frac{\hbar}{2}$ or $-\frac{\hbar}{2}$, a purely quantum mechanical phenomenon. Because its value is always the same, it is regarded as an intrinsic property of electrons, and is sometimes known as “intrinsic angular momentum” (to distinguish it from orbital angular momentum, which can vary and depends on the presence of other particles).

The act of observing (measuring) the momentum along the z direction corresponds to the operator \hat{S}_z , which project the value of the total spin operator \hat{S}^2 along the z axis. The eigenvalues of the projector operator are:

$$\hat{S}_z \phi = \hbar m_s \phi, \quad (7.2)$$

where $m_s = \{-s, +s\} = \left\{-\frac{1}{2}, +\frac{1}{2}\right\}$ is the spin quantum number along the z component. The eigenvalues for the total spin operator \hat{S}^2 —similarly to the angular momentum operator \hat{L}^2 seen in eq. (6.28)—are:

$$\hat{S}^2 \phi = \hbar^2 s(s+1) \phi, \quad (7.3)$$

The initial state of the particles in the Stern-Gerlach experiment is given by the following wave function:

$$\phi = c_1 \phi_{\uparrow} + c_2 \phi_{\downarrow}, \quad (7.4)$$

where $\uparrow = +\frac{\hbar}{2}$, $\downarrow = -\frac{\hbar}{2}$, and the coefficients c_1 and c_2 are complex numbers. In this initial state, spin can point in any direction. The expectation value of the operator \hat{S}_z (the quantity that the Stern-Gerlach experiment measures), can be obtained using eq. (6.22):

$$\begin{aligned} \langle S_z \rangle &= \int \phi^* \hat{S}_z \phi d\mathbf{s} \\ &= +\frac{\hbar}{2} |c_1|^2 - \frac{\hbar}{2} |c_2|^2, \end{aligned} \quad (7.5)$$

where the integration is performed along a special coordinate \mathbf{s} composed of only two values, and the coefficient c_1 and c_2 are complex numbers. Applying the normalization condition, eq. (6.20) we can obtain:

$$|c_1|^2 + |c_2|^2 = 1 \quad \longrightarrow \quad |c_1|^2 = |c_2|^2 = \frac{1}{2}. \quad (7.6)$$

This equation is not sufficient to determine the values of the coefficients since they are complex numbers. Eq. (7.6), however, tells us that the squared magnitudes of the coefficients can be interpreted as probabilities of outcome from the experiment. This is true because their values are obtained from the normalization condition, and the normalization condition guarantees that the system is observed with probability equal to one. Summarizing, since we started with random initial directions, each of the two states, ϕ_{\uparrow} and ϕ_{\downarrow} , will be observed with equal probability of $\frac{1}{2}$.

7.2 Sequential Stern-Gerlach Experiments

An interesting result can be obtain if we link multiple Stern-Gerlach apparatuses into one experiment and we perform the measurement along two orthogonal directions in space. As we showed in the previous section, all particles leaving the first Stern-Gerlach apparatus are in an eigenstate of the \hat{S}_z operator (i.e., their spin is either “up or”down” with respect to the z -direction). We can then take either one of the two resulting beams (for simplicity let’s take the “spin up” output), and perform another spin

measurement on it. If the second measurement is also aligned along the z -direction then only one outcome will be measured, since all particles are already in the “spin up” eigenstate of \hat{S}_z . In other words, the measurement of a particle being in an eigenstate of the corresponding operator leaves the state unchanged.

If, however, we perform the spin measurement along a direction perpendicular to the original z -axis (i.e., the x -axis) then the output will equally distribute among “spin up” or “spin down” in the x -direction, which in order to avoid confusion, we can call “spin left” and “spin right”. Thus, even though we knew the state of the particles beforehand, in this case the measurement resulted in a random spin flip in either of the measurement directions. Mathematically, this property is expressed by the nonvanishing of the commutator of the spin operators:

$$[\hat{S}_z, \hat{S}_x] \neq 0. \quad (7.7)$$

We can finally repeat the measurement a third time, with the magnet aligned along the original z -direction. According to classical physics, after the second apparatus, we would expect to have one beam with characteristic “spin up” and “spin left”, and another with characteristic “spin up” and “spin right”. The outcome of the third measurement along the original z -axis should be one output with characteristic “spin up”, regardless to which beam the magnet is applied (since the “spin down” component should have been “filtered out” by the first experiment, and the “spin left” and “spin right” component should be filtered out by the third magnet). This is not what is observed. The output of the third measurement is—once again—two beams in the z direction, one with “spin up” characteristic and the other with “spin down”.

This experiment shows that spin is not a classical property. The Stern-Gerlach apparatus does not behave as a simple filter, selecting beams with one specific pre-determined characteristic. The second measurement along the x axis destroys the previous determination of the angular momentum in the z direction. This means that this property cannot be measured on two perpendicular directions at the same time.

7.3 Spin Operators

The mathematics of quantum mechanics tell us that \hat{S}_z and \hat{S}_x do not commute. When two operators do not commute, the two measurable quantities

that are associated with them cannot be known at the same time.

In 3-dimensional space there are three directions that are orthogonal to each other $\{x, y, z\}$. Thus, we can define a third spin projection operator along the y direction, \hat{S}_y , corresponding to a new set of Stern-Gerlach experiments where the second magnet is oriented along a direction that is orthogonal to the two that we consider in the previous section. The total spin operator, \hat{S}^2 , can then be constructed similarly to the total angular momentum operator of eq. (6.27), as:

$$\begin{aligned}\hat{S}^2 &= \hat{S} \cdot \hat{S} = (\mathbf{i}\hat{S}_x + \mathbf{j}\hat{S}_y + \mathbf{k}\hat{S}_z) \cdot (\mathbf{i}\hat{S}_x + \mathbf{j}\hat{S}_y + \mathbf{k}\hat{S}_z) \\ &= \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2,\end{aligned}\quad (7.8)$$

with $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ the unitary vectors in three-dimensional space.

Wolfgang Pauli explicitly derived the relationships between all three spin projection operators. Assuming the magnetic field along the z axis, Pauli's relations can be written using simple equations involving the two possible eigenstates ϕ_\uparrow and ϕ_\downarrow :

$$\begin{aligned}\hat{S}_x \phi_\uparrow &= \frac{\hbar}{2} \phi_\downarrow & \hat{S}_y \phi_\uparrow &= \frac{\hbar}{2} i \phi_\downarrow & \hat{S}_z \phi_\uparrow &= \frac{\hbar}{2} \phi_\uparrow \\ \hat{S}_x \phi_\downarrow &= \frac{\hbar}{2} \phi_\uparrow & \hat{S}_y \phi_\downarrow &= -\frac{\hbar}{2} i \phi_\uparrow & \hat{S}_z \phi_\downarrow &= -\frac{\hbar}{2} \phi_\downarrow,\end{aligned}\quad (7.9)$$

where i is the imaginary unit ($i^2 = -1$). In other words, for \hat{S}_z we have eigenvalue equations, while the remaining components have the effect of permuting state ϕ_\uparrow with state ϕ_\downarrow after multiplication by suitable constants. We can use these equations, together with eq. (6.7), to calculate the commutator for each couple of spin projector operators:

$$\begin{aligned}[\hat{S}_x, \hat{S}_y] &= i\hbar \hat{S}_z \\ [\hat{S}_y, \hat{S}_z] &= i\hbar \hat{S}_x \\ [\hat{S}_z, \hat{S}_x] &= i\hbar \hat{S}_y,\end{aligned}\quad (7.10)$$

which prove that the three projection operators do not commute with each other.

Example 7.1. *Proof of Commutator Between Spin Projection Operators.*

The equations in (7.10) can be proved by writing the full eigenvalue equation and solving it using the definition of commutator, eq.(6.7), in conjunction with Pauli's relation, eqs. (7.9). For example, for the first couple:

$$\begin{aligned}
 [\hat{S}_x, \hat{S}_y] \phi_{\uparrow} &= \hat{S}_x \hat{S}_y \phi_{\uparrow} - \hat{S}_y \hat{S}_x \phi_{\uparrow} \\
 &= \hat{S}_x \left(\frac{\hbar}{2} i \phi_{\downarrow} \right) - \hat{S}_y \left(\frac{\hbar}{2} \phi_{\downarrow} \right) \\
 &= \frac{\hbar}{2} \left(\frac{\hbar}{2} i \phi_{\downarrow} \right) - \left(-\frac{\hbar}{2} i \right) \left(\frac{\hbar}{2} \phi_{\downarrow} \right) \\
 &= \left(\frac{\hbar^2}{4} + \frac{\hbar^2}{4} \right) i \phi_{\uparrow} \\
 &= \frac{\hbar^2}{2} i \phi_{\uparrow} \\
 &= i \hbar \hat{S}_z \phi_{\uparrow}
 \end{aligned} \tag{7.11}$$

7.4 Chapter Review

7.4.1 Study Questions

1. The Stern–Gerlach experiment with silver atoms showed that the beam splits into what?
 - a. a continuous smear of intensities
 - b. three discrete spots ($m = -1, 0, +1$)
 - c. a single undivided spot
 - d. two and only two discrete spots on the detector
 - e. four discrete spots
-

2. What key quantum property is demonstrated by the observed splitting in the Stern–Gerlach experiment?

- a. continuous distribution of angular momentum orientations
 - b. quantization of spin angular momentum along the field direction
 - c. conservation of classical orbits
 - d. absence of magnetic moments in atoms
 - e. violation of energy conservation
-

3. For an electron (a spin- $\frac{1}{2}$ particle), what are the possible measured values of the spin component s_z ?

- a. $0, \pm \hbar$
 - b. $\pm \hbar$
 - c. Only $+\hbar/2$
 - d. Any real multiple of \hbar
 - e. $+\hbar/2$ or $-\hbar/2$
-

4. In a Stern–Gerlach experiment measuring s_z , a general spin state can be written as:

- a. A classical vector of length $\hbar/2$
- b. A single eigenstate of \hat{s}_z only
- c. $c_1 |\uparrow\rangle + c_2 |\downarrow\rangle$ with complex c_1, c_2

- d. A mixture of infinitely many s_z eigenvalues
 - e. A real linear combination of position eigenstates
-

5. After a beam passes through a Stern–Gerlach apparatus aligned along z and only the “spin up” channel is allowed through, what is the spin state of the particles that pass through?

- a. An equal superposition of up and down
 - b. Eigenstate $|\downarrow_z\rangle$ of \hat{s}_z
 - c. Eigenstate $|\uparrow_z\rangle$ of \hat{s}_z
 - d. A classical mixture of all directions
 - e. An eigenstate of \hat{s}_x
-

6. If a second Stern–Gerlach apparatus is placed after the first and is also aligned along the same z direction, what happens to a beam that is already in $|\uparrow_z\rangle$?

- a. It is entirely transmitted as $|\uparrow_z\rangle$ with no further splitting
 - b. It becomes completely unpolarized
 - c. It splits again into up and down beams
 - d. It is entirely absorbed
 - e. It flips to $|\downarrow_z\rangle$
-

7. In a sequential Stern–Gerlach experiment, a z -up beam is sent into a second apparatus aligned along x . what is observed at the second apparatus?

- a. Only one output beam (“spin left”)
 - b. Two beams with equal intensities: “spin left” and “spin right” along x
 - c. Only one output beam (“spin right”)
 - d. A continuous spread in deflection angles
 - e. No beam at all
-

8. Mathematically, the non-classical nature of sequential Stern–Gerlach experiments is expressed by which property of spin operators such as \hat{s}_z and \hat{s}_x ?

- a. They commute: $[\hat{s}_z, \hat{s}_x] = 0$
 - b. They have identical eigenvalues
 - c. They are proportional to the identity
 - d. They do not commute: $[\hat{s}_z, \hat{s}_x] \neq 0$
 - e. They vanish when squared
-

9. In three dimensions, the spin components are generally associated with which three operators?

- a. $\hat{s}_r, \hat{s}_y, \hat{s}_z$
- b. $\hat{s}_x, \hat{s}_y, \hat{s}_z$
- c. $\hat{s}_+, \hat{s}_-, \hat{s}_z$
- d. $\hat{s}_1, \hat{s}_2, \hat{s}_3$ with all commuting

e. $\hat{s}_x, \hat{s}_z, \hat{s}_t$

10. What key lesson about spin is emphasized by the Stern–Gerlach experiments?

- a. Spin is a small classical correction to orbital motion
 - b. Spin has a direct classical analogue in rotating spheres
 - c. Spin is a purely quantum property with no classical analogue
 - d. Spin is irrelevant for chemistry
 - e. Spin is always conserved and cannot change under measurement
-

Answers: 1.d, 2.b, 3.e, 4.c, 5.c, 6.a, 7.b, 8.d, 9.b, 10.c

Chapter 8

Postulates of Quantum Mechanics

In order to understand deeper quantum mechanics, scientists have derived a series of axioms that result in what are called *postulates of quantum mechanics*. These are, in fact, assumptions that we need to make to understand how the measured reality relates with the mathematics of quantum mechanics. It is important to notice that the postulates are necessary for the interpretation of the theory, but not for the mathematics behind it. Regarding of whether we interpret it or not, the mathematics is complete and consistent. In fact, as we will see in the next chapter, several controversies regarding the interpretation of the mathematics are still open, and different philosophies have been developed to rationalize the results. Recall also that there are different ways of writing the equation of quantum mechanics, all equivalent to each other (i.e., Schrödinger's differential formulation and Heisenberg's algebraic formulation that we saw in chapter 3). For these reasons, there is not an agreement on the number of postulates that are necessary to interpret the theory, and some philosophy and/or formulation might require more postulates than others. In this chapter, we will discuss the six postulates, as they are usually presented in chemistry and introductory physics textbooks and as they relate with a basic statistical interpretation of quantum mechanics. Regardless of the philosophical consideration on the meanings and numbers of the postulate, as well as their physical origin, these statements will make the interpretation of the theory a little easier, as we will see in the next chapter.

8.1 Postulate 1: The Wave Function Postulate

The state of a quantum mechanical system is completely specified by a function $\Psi(\mathbf{r}, t)$ that depends on the coordinates of the particle(s) and on time. This function, called the wave function or state function, has the important property that $\Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t)d\tau$ is the probability that the particle lies in the volume element $d\tau$ located at \mathbf{r} at time t .

The wave function must satisfy certain mathematical conditions because of this probabilistic interpretation. For the case of a single particle, the probability of finding it somewhere is 1, so that we have the normalization condition

$$\int_{-\infty}^{\infty} \Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t)d\tau = 1 \quad (8.1)$$

It is customary to also normalize many-particle wave functions to 1. As we already saw for the particle in a box in chapter 4, a consequence of the first postulate is that the wave function must also be single-valued, continuous, and finite, so that derivatives can be defined and calculated at each point in space. This consequence allows for operators (which typically involve derivation) to be applied without mathematical issues.

8.2 Postulate 2: Experimental Observables

To every observable in classical mechanics there corresponds a linear, Hermitian operator in quantum mechanics. We have in part already discussed this postulate in chapter 6, albeit we didn't call it as such. This postulate is necessary if we require the expectation value of an operator \hat{A} to be real, as it should be.

8.3 Postulate 3: Individual Measurements

In any measurement of the observable associated with operator \hat{A} , the only values that will ever be observed are the eigenvalues a that satisfy the eigenvalue equation:

$$\hat{A}\Psi = a\Psi. \quad (8.2)$$

This postulate captures the central point of quantum mechanics: the values of dynamical variables can be quantized (although it is still possible to have a continuum of eigenvalues in the case of unbound states). If the system is in an eigenstate of \hat{A} with eigenvalue a , then any measurement of the quantity A will yield a . Although measurements must always yield an eigenvalue, the state does not have to be an eigenstate of \hat{A} initially.

An arbitrary state can be expanded in the complete set of eigenvectors of \hat{A} ($\hat{A}\Psi_i = a_i\Psi_i$) as:

$$\Psi = \sum_i^n c_i \Psi_i, \quad (8.3)$$

where n may go to infinity. In this case, we only know that the measurement of A will yield one of the values a_i , but we don't know which one. However, we do know the probability that eigenvalue a_i will occur (it is the absolute value squared of the coefficient, $|c_i|^2$, as we obtained already in chapter 6), leading to the fourth postulate below.

8.4 Postulate 4: Expectation Values and Collapse of the Wavefunction

If a system is in a state described by a normalized wave function Ψ , then the average value of the observable corresponding to \hat{A} is given by:

$$\langle A \rangle = \int_{-\infty}^{\infty} \Psi^* \hat{A} \Psi d\tau. \quad (8.4)$$

An important consequence of the fourth postulate is that, after measurement of Ψ yields some eigenvalue a_i , the wave function immediately “collapses” into the corresponding eigenstate Ψ_i . In other words, measurement affects the state of the system. This fact is used in many experimental tests of quantum mechanics, such as the Stern-Gerlach experiment. Think again at the sequential experiment that we discussed in chapter 7. The act of measuring the spin along one coordinate is not simply a “filtration” of some pre-existing feature of the wave function, but rather an act that changes the nature of the wave function itself, affecting the outcome of future experiments. To this act corresponds the collapse of the wave function, a process that remains unexplained to date. Notice how the controversy is not in the

mathematics of the experiment, which we already discussed in the previous chapter without issues. The issues rather arise because we don't know how to define the measurement act in itself (other than the fact that it is some form of quantum mechanical procedure with clear and well-defined macroscopic outcomes). This is the reason why the collapse of the wave function is also sometimes called the *measurement problem* of quantum mechanics, and it is still a source of research and debate among modern scientists.

8.5 Postulate 5: Time Evolution

The wave function of a system evolves in time according to the time-dependent Schrödinger equation:

$$\hat{H}\Psi(\mathbf{r}, t) = i\hbar \frac{\partial\Psi}{\partial t}. \quad (8.5)$$

The central equation of quantum mechanics must be accepted as a postulate.

8.6 Postulate 6: Pauli Exclusion Principle

If we want to describe a system with N particles, we need to account not only for the position of each particle, but also for their spin. This happens because the quantity that is physically observable in quantum mechanics is not the wave function itself, but rather its square. We can start to understand why this happens by writing a multi-particle wave function using $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$, with $\mathbf{r}_1 = \{x_1, y_1, z_1, s_1\}$ being the quantum state of particle 1 composed of its position in real space $\{x, y, z\}$ plus its spin state s , $\mathbf{r}_2 = \{x_2, y_2, z_2, s_2\}$ the quantum state of particle 2, and so on. The physically observable quantity for such a wave function will be $|\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2$. When we exchange the quantum state of two particles (let's say 1 and 2), this observable quantity should remain the same, since quantum particles are indistinguishable from each other:

$$|\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2 = |\Psi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N)|^2. \quad (8.6)$$

This equality, however, can be fulfilled in two different ways:

$$\begin{aligned} \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) &= -\Psi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N) && \text{fermions,} \\ \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) &= +\Psi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N) && \text{bosons.} \end{aligned} \quad (8.7)$$

Each of the two options above characterize a different kind of particles. Those that have an antisymmetric wave function with respect to exchange of quantum states are called **fermions**, while those that have a symmetric one are called **bosons**. Eq. (8.7) is known in physics as *the spin-statistics theorem*.

But why would the spin of a particle be important in the description of a multi-particle system? To understand this fact we can use the spin-statistics theorem to derive the much more popular¹ *Pauli exclusion principle*. Let's start with a multi-fermionic wave function:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = -\Psi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N), \quad (8.8)$$

and then let's force two fermions (for simplicity of notation, let's take 1 and 2) into the same state, $\mathbf{r}_1 = \mathbf{r}_2$. So if we exchange the quantum states of these two fermions, nothing can change, because they are in the same state (we forced them to be in it), therefore we must also have:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \Psi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N). \quad (8.9)$$

The only way both eqs. (8.8) and (8.9) can be fulfilled simultaneously is if $\Psi = 0$. From this we can derive the Pauli exclusion principle, that says that if we have two fermions, they cannot share the same quantum state (the wave function for such case will need to be exactly zero everywhere). In summary, if we need to describe a system composed of many fermions, then we need to account for all of their individual spin. That is because if two fermions share the same position in space $\{x, y, z\}$, then their spin must be different. Or alternatively, if two fermions have the same spin, they will never occupy the same position in space. Notice that using a similar argument it's very easy to demonstrate that there is no Pauli exclusion principle for Bosons. Several bosons can all occupy the same position in space, regardless of their spin.² Unfortunately (or perhaps fortunately), electrons are bosons...

¹at least in chemistry

²See, for example, Bose-Einstein condensates here.

8.7 Chapter Review

8.7.1 Study Questions

1. According to postulate 1 in this chapter, what completely specifies the state of a quantum system?

- a. A set of classical coordinates and momenta
 - b. A normalized wave function $\psi(\mathbf{r}, t)$ over configuration space
 - c. The list of all possible measurement outcomes
 - d. The expectation values of all observables
 - e. The energy and total angular momentum
-

2. In the wave function postulate, $|\psi(\mathbf{r}, t)|^2 d\tau$ is interpreted as:

- a. The probability that the particle is found in volume element $d\tau$ around \mathbf{r} at time t
 - b. The charge density of the particle
 - c. The kinetic energy density
 - d. The number of particles per unit volume
 - e. The potential energy at \mathbf{r}
-

3. Postulate 2 (experimental observables) states that to every classical observable there corresponds what?

- a. A real-valued function of time only
- b. A complex number

- c. A linear Hermitian operator acting on wave functions
 - d. A non-linear operator that may not be hermitian
 - e. A classical random variable
-

4. Why must operators representing physical observables be Hermitian?

- a. To guarantee that eigenvalues are integers
 - b. To ensure that eigenfunctions are always real
 - c. To guarantee real measurement outcomes (real eigenvalues)
 - d. To make the Hamiltonian time independent
 - e. To enforce energy conservation in all processes
-

5. If the system is in an eigenstate of \hat{a} with eigenvalue a_k , what does postulate 3 say about the outcome of a measurement of a ?

- a. The outcome is random among all eigenvalues
 - b. The outcome is always a_k
 - c. The outcome is always $\langle a \rangle$
 - d. The outcome is one of two possible eigenvalues with equal probability
 - e. The outcome is complex
-

6. In postulate 4, once a measurement yields a particular eigenvalue a_k , what happens to the wave function (in the idealized description)?

- a. It remains unchanged
 - b. It collapses to the corresponding eigenstate of \hat{a}
 - c. It becomes identically zero
 - d. It becomes a uniform superposition of all eigenstates
 - e. It collapses to the ground state of the hamiltonian
-

7. For a system in a superposition $\psi = \sum_k c_k \psi_k$ of \hat{a} -eigenfunctions ψ_k , what is the probability of obtaining eigenvalue a_j upon measuring a ?

- a. $|c_j|^2$ (assuming normalized eigenfunctions)
 - b. $|a_j|^2$
 - c. $1/|c_j|^2$
 - d. Always $1/n$, where n is the number of eigenvalues
 - e. $\text{Re}(c_j)$
-

8. Postulate 5 implies that the time-evolution operator must have which property?

- a. It is unitary, preserving normalization of ψ
- b. It is Hermitian
- c. It is real and symmetric
- d. It is diagonal in the position representation
- e. It always commutes with position

9. Postulate 6 in this chapter (Pauli exclusion principle) applies to which type of particles?

- a. All particles, regardless of spin
 - b. Bosons only
 - c. Fermions with half-integer spin (e.g., electrons)
 - d. Photons only
 - e. Classical point particles
-

10. What does the Pauli exclusion principle state about the many-electron wave function for identical fermions?

- a. It must be symmetric under exchange of any two identical fermions
 - b. It must vanish everywhere
 - c. It must be antisymmetric under exchange of any two identical fermions
 - d. It must be purely real
 - e. It must be constant in space
-

Answers: 1.b, 2.a, 3.c, 4.c, 5.b, 6.b, 7.a, 8.a, 9.c, 10.c

Chapter 9

Quantum Weirdness

In this chapter, we will delve deeper into the strangeness of quantum mechanics. In particular, we will explore quantum phenomena that don't have a classical counterpart, starting from perhaps the most simple but also one of the most revealing: the double-slit experiment.

9.1 The double-slit experiment

The double-slit experiment is considered by many the seminal experiment in quantum mechanics. The reason why we see it only at this advanced point is that its interpretation is not as straightforward as it might seem from a superficial analysis. The famous physicist Richard Feynman was so fond of this experiment that he used to say that all of quantum mechanics can be understood from carefully thinking through its implications.

The premises of the experiment are very simple: cut two slits in a solid material (such as a sheet of metal), send light or electrons through them, and observe what happens on a screen position at some distance on the other side. The result of this experiment though are far from straightforward.

Let's first consider the single-slit case. If light consisted of classical particles, and these particles were sent in a straight line through a single-slit and allowed to strike a screen on the other side, we would expect to see a pattern corresponding to the size and shape of the slit. However, when this "single-slit experiment" is actually performed, the pattern on the screen is a diffraction pattern in which the light is spread out. The smaller the slit,

the greater the angle of spread. This behavior is typical of waves, where diffraction explains the pattern as being the result of the interference of the waves with the slit.

If one illuminates two parallel slits, the light from the two slits again interferes. Here the interference is a more pronounced pattern with a series of alternating light and dark bands. The width of the bands is a property of the frequency of the illuminating light. The pattern observed on the screen is the result of this interference, as shown in figure 9.1.¹

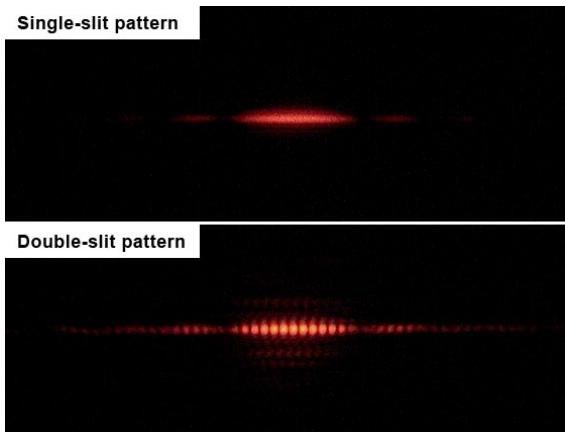


Figure 9.1: Outcomes of single-slit and double-slit experiments.

The interference pattern resulting from the double-slit experiment are observed not only with light, but also with a beam of electrons, and other small particles.

9.1.1 The individual particles experiment

The first twist in the plot is if we perform the experiment by sending individual particles (e.g. either individual photons, or individual electrons). Sending particles through a double-slit apparatus one at a time results in single particles appearing on the screen, as expected. Remarkably, however, an interference pattern emerges when these particles are allowed to build up

¹This diagram is taken from Wikipedia by user Jordgette, and distributed under CC BY-SA 3.0 license.

one by one (figure 9.2²). The resulting pattern on the screen is the same as if each individual particle had passed through both slits.

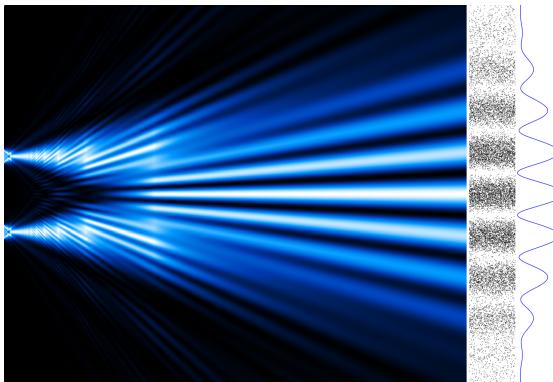


Figure 9.2: Numerical simulation of the double-slit experiment with electrons.

This variation of the double-slit experiment demonstrates the wave–particle duality: the particle is measured as a single pulse at a single position, while the wave describes the probability of absorbing the particle at a specific place on the screen.

9.1.2 “Which way” experiment

A second twist happens if we place particle detectors at the slits with the intent of showing through which slit a particle goes. The interference pattern in this case will disappear.

This experiment illustrates that photons (and electrons) can behave as either particles or waves, but cannot be observed as both at the same time. The simplest interpretation of this experiment is that the wave function of the photon collapses into a deterministic position due to the interaction with the detector on the slit, and the interference pattern is therefore lost. This result also proves that in order to measure (detect) a photon, we must interact with it, an act that changes its wave function.

²This diagram is taken from Wikipedia by user Alexandre Gondran, and distributed under CC BY-SA 4.0 license

The interpretation of the results of this experiment is not simple. As for other situations in quantum mechanics, the problem arise not because we cannot describe the experiment in mathematical terms, but because the math that we need to describe it cannot be related to the macroscopic classical world we live in. According to the math, in fact, particles in the experiment are described exclusively in probabilistic terms (given by the square of the wave function). The macroscopic world, however, is not probabilistic, and outcomes of experiments can be univocally measured. Several different ways of redeeming this controversy have been proposed, including for example the possibility that quantum mechanics is incomplete (the emergence of probability is due to the ignorance of some more fundamental deterministic feature of nature), or assuming that every time a measurement is done on a quantum system, the universe splits, and every possible measurable outcome is observed in different branches of our universe (we only happen to live in one of such branches, so we observe only one non-probabilistic result).³ The interpretation of quantum mechanics is still an unsolved problem in modern physics (luckily, it does not prevent us from using quantum mechanics in chemistry).

9.2 Heisenberg's Uncertainty Principle

Let's now revisit the simple case of a free particle. As we saw in chapter 4, the wave function that solved the TISEq:

$$\psi(x) = A \exp(\pm ikx), \quad (9.1)$$

is the equation of a plane wave along the x direction. This result is in agreement with the de Broglie hypothesis, which says that every object in the universe is a wave. If this wave function describes a particle with mass (such as an electron), freely moving along one spatial direction x , it would be reasonable to ask the question: where is the particle located? Analyzing eq. (9.1), however, it is not possible to answer this question since $\psi(x)$ is delocalized in space from $x = -\infty$ to $x = +\infty$.⁴ In other words, the particle position is extremely uncertain because it could be essentially anywhere along the wave.

³The interested student can read more about different interpretations [HERE](#).

⁴The time-dependent picture does not help us either, but since it is a little more complicated to work with the TDSEq, we are not showing it here.

Thus for a free particle, the particle side of the wave-particle duality seems completely lost. We can, however, bring it back into the picture by writing the wave function as a sum of many plane waves, called a *wave packet*:

$$\psi(x) \propto \sum_n A_n \exp\left(\frac{ip_n x}{\hbar}\right), \quad (9.2)$$

where A_n represents the relative contribution of the mode p_n to the overall total. We are allowed to write the wave function this way because each individual plane wave is a solution of the TISEq, and as we already saw in chapter 6 and several other places, the sum of each individual solution is also a solution. An interesting consequence of writing the wave function as a wave packet is that when we sum different waves, they interfere with each other, and they might localize in some region of space. Thus for a wave function written as in eq. (9.2), the wave packet can become more localized. We may also make this procedure a step further to the continuum limit, where the wave function goes from a sum to an integral over all possible modes:

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \varphi(p) \cdot \exp\left(\frac{ipx}{\hbar}\right) dp, \quad (9.3)$$

where $\varphi(p)$ represents the amplitude of these modes and is called the wave function in momentum space. In mathematical terms, we say that $\varphi(p)$ is the Fourier transform of $\psi(x)$ and that x and p are conjugate variables. Adding together all of these plane waves comes at a cost; namely, the momentum has become less precise since it becomes a mixture of waves of many different momenta.

One way to quantify the precision of the position and momentum is the standard deviation, σ . Since $|\psi(x)|^2$ is a probability density function for position, we calculate its standard deviation. The precision of the position is improved—i.e., reduced σ_x —by using many plane waves, thereby weakening the precision of the momentum—i.e., increased σ_p . Another way of stating this is that σ_x and σ_p have an inverse relationship (once we know one with absolute precision, the other becomes completely unknown). This fact was discovered by Werner Heisenberg and is now called the **Heisenberg's uncertainty principle**. The mathematical treatment of this procedure results in the simple formula:

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}. \quad (9.4)$$

The uncertainty principle can be extended to any couple of conjugated variables, including, for example, energy and time, angular momentum components along perpendicular directions, spin components along perpendicular directions, etc. It is also easy to show that conjugate variables in quantum mechanics correspond to non-commuting operators.⁵

9.3 Tunneling

Tunneling is a phenomenon where a particle may cross a barrier even if it does not have sufficient kinetic energy to overcome the potential of the barrier itself. In this situation, the particle is said to “tunnel through” the barrier following a purely quantum mechanical phenomenon (figure 9.3).⁶

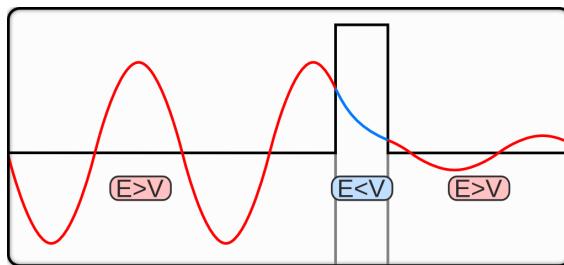


Figure 9.3: Quantum tunneling through a barrier. The energy of the tunneled particle is the same but the probability amplitude is decreased.

To explain tunneling we must resort once again to the TISeq. A traveling or standing wave function incident on a non-infinite potential barrier (V_0) decays in the potential as a function of $A_0 \exp[-\alpha x]$, where A_0 is the amplitude at the boundary, α is proportional to the potential, and x is the distance into the potential. If a second well exists at infinite distance from the first well, the probability goes to zero, so the probability of a particle existing in the second well is zero. If a second well is brought closer to the first well, the amplitude of the wave function at this boundary is not zero, so the particle may tunnel into that well from the first well. It would appear that the particle is “leaking” through the barrier; it can travel through it

⁵Therefore, a simpler way of finding if two variables are subject to the uncertainty principle is to check if their corresponding operators commute.

⁶This diagram is taken from Wikipedia by user Felix Kling, and distributed under CC BY-SA 3.0 license.

without having to surmount it. An important point to keep in mind is that tunneling conserves energy. The final sum of the kinetic and potential energy of the system cannot exceed the initial sum. Therefore, the potential on both sides of the barrier does not need to be the same, but the sum of the ground state energy and the potential on the opposite side of the barrier may not be larger than the initial particle energy and potential.

Tunneling can be described using the TISEq, eq. (6.23). For the tunneling problem we can take the potential V to be zero for all space, except for the region inside the barrier (between 0 and l):

$$V = \begin{cases} 0 & \text{if } -\infty < x \leq 0 \\ V_0 & \text{if } 0 < x < l \\ 0 & \text{if } l \leq x < \infty \end{cases}. \quad (9.5)$$

To solve the TISEq with this potential, we must solve it separately for each region, but we should make sure that the wave function stays single-valued, continuous and everywhere continuously differentiable. The general solution for each region, before applying the boundary conditions, is:

$$\psi = \begin{cases} A \sin(kx) + B \cos(kx) & \text{if } -\infty < x \leq 0 \\ C \exp(-\alpha x) + D \exp(\alpha x) & \text{if } 0 < x < l \\ E \sin(kx) + F \cos(kx) & \text{if } l \leq x < \infty \end{cases} \quad (9.6)$$

where $k = \frac{\sqrt{2mE}}{\hbar}$, and $\alpha = \frac{\sqrt{2m(V_0-E)}}{\hbar}$. To enforce continuity, we must have at the first boundary:

$$A \sin(0) + B \cos(0) = C \exp(0) + D \exp(0), \quad (9.7)$$

which implies that $A = 0$, and $B = C + D$. At the opposite boundary:

$$A \sin(kl) + B \cos(kl) = C \exp(-\alpha l) + D \exp(\alpha l). \quad (9.8)$$

We notice that, as l goes to infinity, the right hand side of eq. (9.8) goes to infinity, which does not make physical sense. To reconcile this, we must set $D = 0$.

For the final region, E and F , present a potentially intractable problem. However, if one realizes that the value at the boundary l is driving the wave

in the region l to infinity, it may also be realized that the wave function could be rewritten as $C \exp[-\alpha l] \cos[k(x - l)]$, phase shifting the wave function by the value of l , and setting the amplitude to the boundary value. Summarizing, the wave function is:

$$\psi = \begin{cases} B \cos(kx) & \text{if } -\infty < x \leq 0 \\ B \exp(-\alpha x) & \text{if } 0 < x < l \\ B \exp(-\alpha l) \cos[k(x - l)] & \text{if } l \leq x < \infty. \end{cases} \quad (9.9)$$

Comparing the wave function on the left of the barrier with the one on its right, we notice how the amplitude is attenuated by the barrier as $\exp\left(-l \frac{\sqrt{2m(V_0 - E)}}{\hbar}\right)$, where l is the width of the barrier, and $(V_0 - E)$ is the difference between the potential energy of the barrier and the current energy of the particle. Since the square of the wave function is the probability distribution, the probability of transmission through a barrier is:

$$\exp\left(-2l \frac{\sqrt{2m(V_0 - E)}}{\hbar}\right). \quad (9.10)$$

As the barrier width or height approaches zero, the probability of a particle tunneling through the barrier becomes one. We can also note that k is unchanged on the other side of the barrier. This implies that the energy of the particle is exactly the same as it was before it tunneled through the barrier, as stated earlier, the only thing that changes is the quantity of particles going in that direction. The rest is reflected off the barrier, and go back the way it came. On the opposite end, as the barrier width or height approaches infinity, the probability of a particle tunneling through the barrier becomes zero, and the barrier behaves similarly to those that contained the particle in the particle in a box example discussed in chapter 4.

9.3.1 Tunneling in chemical reactions

The ammonia umbrella inversion reaction provides an excellent example of the impact of quantum mechanical tunneling on the kinetic isotope effect (KIE). The reaction involves the inversion of the nitrogen atom in ammonia, which changes its position from above the plane of the three hydrogen atoms to below it and vice versa. This is a rapid oscillation of the atom

and substituents, classically described by the molecule passing through a planar transition state. The ammonia interconversion is, however, more rapid than the classical prediction would predict. At room temperature, ammonia inverts 30 billion times per second. Two factors contribute to the rapidity of the inversion: a low energy barrier (24.2 kJ/mol) and a narrow width of the barrier itself, which allows for frequent quantum tunneling of the hydrogen atoms. In contrast, phosphine (PH_3) inverts very slowly at room temperature (energy barrier: 132 kJ/mol).

One of the primary effects of tunneling can be seen on the KIE. This is because tunneling allows some of the isotopes to bypass the activation barrier, making it easier for them to react. As a result, the reaction rate for the heavier isotope is increased relative to what it would be in the absence of tunneling, leading to a reduction in the KIE. In the case of the umbrella inversion reaction, replacing ammonia with its deuterated counterpart, ND_3 , leads to a higher activation barrier, resulting in a slower reaction rate.

The impact of tunneling on the KIE can be quantified using transition state theory (TST), which accounts for the probability of tunneling through the activation barrier. The KIE is given by:

$$\text{KIE} = \frac{k(\text{H})}{k(\text{D})} = \left[\frac{m(\text{H})}{m(\text{D})} \right]^{\frac{1}{2}} \frac{k_{\text{tunneling}}}{k_{\text{classical}}}, \quad (9.11)$$

where $k(\text{H})$ and $k(\text{D})$ are the rate constants for the reaction involving hydrogen and deuterium, respectively; $m(\text{H})$ and $m(\text{D})$ are the masses of hydrogen and deuterium; $k_{\text{tunneling}}$ is the rate constant for tunneling; and $k_{\text{classical}}$ is the rate constant in the absence of tunneling.

Using experimental data for the umbrella inversion reaction, the KIE can be calculated to be approximately 1.3 in the absence of tunneling. However, when tunneling is allowed, the KIE is reduced to around 1.1. This reduction in the KIE due to tunneling is a significant effect and must be considered when studying chemical reactions involving light atoms.

9.4 Chapter Review

9.4.1 Study Questions

1. In the double-slit experiment with single particles (e.g., electrons), what is observed on the screen after many particles have

passed?

- a. a single sharp peak at the center
 - b. two non-overlapping bands corresponding to the two slits
 - c. a classical shadow pattern with uniform intensity between slits
 - d. an interference pattern with alternating bright and dark fringes
 - e. a completely uniform distribution over the screen
-

2. When detectors are placed at the slits to determine “which slit” each particle goes through, what typically happens to the interference pattern?

- a. it becomes sharper and more pronounced
 - b. it disappears, leaving two classical-like bands
 - c. it reverses (bright and dark fringes swap)
 - d. it doubles in fringe number
 - e. it becomes independent of slit separation
-

3. In the Heisenberg uncertainty principle for position and momentum along x , eq. (9.4), what do σ_x and σ_p represent?

- a. exact measurement errors of the apparatus
- b. spreads (standard deviations) in the probability distributions of x and p_x
- c. mean values of position and momentum

- d. classical trajectory deviations
 - e. arbitrary experimental tolerances chosen by the user
-

4. which pair of observables is *not* a standard example of canonically conjugate variables in the context of uncertainty relations?

- a. position and momentum
 - b. energy and time (in a suitable sense)
 - c. different components of spin
 - d. particle number and phase (in some systems)
 - e. mass and electric charge
-

5. Quantum tunneling refers to the phenomenon where a particle does what?

- a. moves faster than light through a barrier
 - b. passes through a region where its classical kinetic energy would be negative
 - c. reflects perfectly from any finite barrier
 - d. ceases to exist inside a barrier
 - e. behaves as a classical wave inside a barrier but as a particle outside
-

6. In a 1d potential-barrier problem, what happens to the wave function in a classically forbidden region where $E < V(x)$?

- a. it vanishes identically
 - b. it oscillates with constant amplitude
 - c. it decays (or grows) exponentially with distance into the barrier
 - d. it becomes a delta function at the barrier edge
 - e. it becomes purely imaginary and non-normalizable
-

7. For a finite barrier of width l , which statement about transmission probability in the tunneling example is qualitatively correct?

- a. it is exactly zero if $E < V_0$
 - b. it is $1/2$ regardless of barrier parameters
 - c. it decreases rapidly as barrier width and/or height increase
 - d. it is independent of the particle's mass
 - e. it is always 1 for normalized wave functions
-

8. Which of the following best explains *why* tunneling can occur even when $E < V_0$ for a barrier?

- a. energy is not conserved in quantum mechanics
- b. the wave function penetrates the classically forbidden region and has nonzero amplitude beyond it
- c. the particle temporarily borrows energy from the vacuum
- d. relativistic corrections allow the particle to accelerate
- e. the barrier momentarily disappears due to fluctuations

9. In a chemical reaction coordinate picture, what does tunneling correspond to?

- a. the reactant climbing *over* the transition-state barrier
 - b. the reactant remaining in a metastable well forever
 - c. the system passing *through* the barrier region where classical energy would be insufficient
 - d. the system skipping the transition state entirely with zero probability
 - e. the system following a purely classical minimum-energy path
-

10. Which experimental signature is commonly taken as evidence for significant hydrogen tunneling in reactions?

- a. linear arrhenius plots with classical slopes
 - b. very small kinetic isotope effects ($k_h/k_d \approx 1$ at all temperatures)
 - c. large kinetic isotope effects and curved (non-linear) arrhenius plots
 - d. complete absence of isotope effects
 - e. rate constants independent of temperature and isotope
-

Answers: 1.d, 2.b, 3.b, 4.e, 5.b, 6.c, 7.c, 8.b, 9.c, 10.c

Chapter 10

Many-Electron Atoms

When two or more electrons are present in a system, the TISEq equation cannot be solved analytically. Thus for the vast majority of chemical applications, we must rely on approximate methods. We will explore some of these approximations in this and further chapter, starting from the many-electron atoms (all atoms other than hydrogen). It is important to stress that because of the nature of approximations, this is still a very active field of scientific research, and improved methods are developed every year.

The electronic Hamiltonian for a many-electron atom can be written as:

$$\hat{H}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \sum_{i=1}^N \left(-\frac{\hbar^2}{2m_e} \nabla_i^2 - \frac{Ze^2}{4\pi\epsilon_0 r_i} \right) + \frac{e^2}{4\pi\epsilon_0} \sum_{i < j} \frac{1}{r_{ij}}, \quad (10.1)$$

where Z is the nuclear charge, m_e and e are respectively the mass and charge of an electron, \mathbf{r}_i and ∇_i^2 are the spatial coordinates and the Laplacian of each electron, $r_i = |\mathbf{r}_i|$, and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the distance between two electrons (all other symbols have been explained in previous chapters). The TISEq is easily written using eq. (6.24).

10.1 Many-Electron Wave Functions

When we have more than one electron, the sixth postulate that we discussed in chapter 8 comes into place. In other words, we need to account for the

spin of the electrons and we need the wave function to be antisymmetric with respect to exchange of the coordinates of any two electrons. In order to do so, we can define a new variable \mathbf{x} which represents the set of all four coordinates associated with an electron: three spatial coordinates \mathbf{r} , and one spin coordinate \mathbf{s} , i.e., $\mathbf{x} = \{\mathbf{r}, \mathbf{s}\}$. We can then write the electronic wave function as $\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$, and we require the sixth postulate to hold by writing:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = -\Psi(\mathbf{x}_2, \mathbf{x}_1, \dots, \mathbf{x}_N) \quad (10.2)$$

A very important step in simplifying $\Psi(\mathbf{x})$ is to expand it in terms of a set of one-electron functions. Since we need to take into account the spin coordinate as well, we can define a new function, called **spin-orbital**, by multiplying a spatial orbital by one of the two spin functions:

$$\begin{aligned} \chi(\mathbf{x}) &= \psi(\mathbf{r})\phi_{\uparrow}(\mathbf{s}), \\ \chi(\mathbf{x}) &= \psi(\mathbf{r})\phi_{\downarrow}(\mathbf{s}). \end{aligned} \quad (10.3)$$

Notice that for a given spatial orbital $\psi(\mathbf{r})$, we can form two spin orbitals, one with \uparrow spin, and one with \downarrow spin (since the spin coordinate \mathbf{s} has only two possible values, as already discussed in chapter 7). For the spatial orbitals we can use the same one-particle functions that solve the TISEq for the hydrogen atom, $\psi_{n\ell m_\ell}(\mathbf{r})$ (eq. (5.7) in chapter 5). Notice how each spin-orbital now depends on four quantum numbers, the three for the spatial part, n, ℓ, m_ℓ , plus the spin quantum number m_s . We need to keep in mind, however, that the spin-orbitals, $\chi_{n\ell m_\ell m_s}$, are *not* analytic solutions to the TISEq, so the resulting wave function is *not* the exact wave function of the system, but just an approximation.

Once we have defined one-electron spin-orbitals for each electron in the system, we can use them as the basis for our many-electron wave function. While doing so, we need to make sure to enforce the antisymmetry property of the overall wave function. We will start from the simplest case of an atom with two electrons with coordinates \mathbf{x}_1 and \mathbf{x}_2 , which we put in two spin-orbitals χ_1 and χ_2 . We can write the total wave function as a linear combination of the two spin-orbitals as:

$$\begin{aligned} \Psi(\mathbf{x}_1, \mathbf{x}_2) &= b_{11}\chi_1(\mathbf{x}_1)\chi_1(\mathbf{x}_2) + b_{12}\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2) + \\ &\quad b_{21}\chi_2(\mathbf{x}_1)\chi_1(\mathbf{x}_2) + b_{22}\chi_2(\mathbf{x}_1)\chi_2(\mathbf{x}_2). \end{aligned} \quad (10.4)$$

We then notice that in order for the antisymmetry principle to be obeyed, we need $b_{12} = -b_{21}$ and $b_{11} = b_{22} = 0$, which give:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = b_{12} [\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2) - \chi_2(\mathbf{x}_1)\chi_1(\mathbf{x}_2)]. \quad (10.5)$$

This wave function is sufficient to describe two-electron atoms and ions, such as helium. The numerical coefficient can be determined imposing the normalization condition, and is equal to $b_{12} = \frac{1}{\sqrt{2}}$. For the ground state of helium, we can replace the spatial component of each spin-orbital with the $1s$ hydrogenic orbital, ψ_{100} , resulting in:

$$\begin{aligned} \Psi(\mathbf{x}_1, \mathbf{x}_2) &= \frac{1}{\sqrt{2}} [\psi_{100}(\mathbf{r}_1)\phi_{\uparrow} \psi_{100}(\mathbf{r}_2)\phi_{\downarrow} - \psi_{100}(\mathbf{r}_1)\phi_{\downarrow} \psi_{100}(\mathbf{r}_2)\phi_{\uparrow}] \\ &= \psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2) \frac{1}{\sqrt{2}} [\phi_{\uparrow}\phi_{\downarrow} - \phi_{\downarrow}\phi_{\uparrow}], \end{aligned} \quad (10.6)$$

which clearly shows how we need just one spatial orbital, ψ_{100} , to describe the system, while the antisymmetry is taken care by a suitable combination of spin functions, $\frac{1}{\sqrt{2}} [\phi_{\uparrow}\phi_{\downarrow} - \phi_{\downarrow}\phi_{\uparrow}]$. Notice also that we commit a small inaccuracy when we say: “two electron occupies one spin-orbital, one electron has spin up, and the other electron has spin down, with configuration: $[\uparrow\downarrow]$ ”, as is typically found in general chemistry textbooks. The reality of the spin configuration is indeed more complicated, and the ground state of helium should be represented as $\frac{1}{\sqrt{2}} [\uparrow\downarrow - \downarrow\uparrow]$.

In order to generalize from two electrons to N , we can first observe how eq. (10.5) could be easily constructed by placing the spin-orbitals into a 2×2 matrix and calculating its determinant:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) \end{vmatrix}, \quad (10.7)$$

where each column contains one spin-orbital, each row contains the coordinates of a single electron, and the vertical bars around the matrix mean that we need to calculate its determinant. This notation is called the **Slater determinant**, and it is the preferred way of building any N -electron wave function. Slater determinants are useful because they can be easily built for any case of N electrons in N spin-orbitals, and they also automatically

enforce the antisymmetry of the resulting wave function. A general Slater determinant is written:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} = |\chi_1, \chi_2, \dots, \chi_N\rangle, \quad (10.8)$$

where the notation $|\cdots\rangle$ is a shorthand to indicate the Slater determinant where only the diagonal elements are reported.

10.2 Approximated Hamiltonians

In order to solve the TISEq for a many-electron atom we also need to approximate the Hamiltonian, since analytic solution using the full Hamiltonian as in eq. (10.1) are impossible to find. The most significant approximation used in chemistry is called the *variational method*.

10.2.1 Variational method

The basic idea of the variational method is to guess a “trial” wave function for the problem consisting of some adjustable parameters called “variational parameters”. These parameters are adjusted until the energy of the trial wave function is minimized. The resulting trial wave function and its corresponding energy are variational method approximations to the exact wave function and energy.

Why would it make sense that the best approximate trial wave function is the one with the lowest energy? This results from the Variational Theorem, which states that the energy of any trial wave function E is always an upper bound to the exact ground state energy \mathcal{E}_0 . This can be proven easily. Let the trial wave function be denoted Φ . Any trial function can formally be expanded as a linear combination of the exact eigenfunctions Ψ_i . Of course, in practice, we don’t know the Ψ_i , since we are applying the variational method to a problem we can’t solve analytically. Nevertheless, that doesn’t

prevent us from using the exact eigenfunctions in our proof, since they certainly exist and form a complete set, even if we don't happen to know them. So, the trial wave function can be written:

$$\Phi = \sum_i c_i \Psi_i, \quad (10.9)$$

and the approximate energy corresponding to this wave function is:

$$E[\Phi] = \frac{\int \Phi^* \hat{H} \Phi d\tau}{\int \Phi^* \Phi d\tau}, \quad (10.10)$$

where $\tau = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ is the ensemble of the spatial coordinates of each electron and the integral symbol is assumed as a $3N$ -dimensional integration. Replacing the expansion over the exact wave functions, we obtain:

$$E[\Phi] = \frac{\sum_{ij} c_i^* c_j \int \Psi_i^* \hat{H} \Psi_j d\tau}{\sum_{ij} c_i^* c_j \int \Psi_i^* \Psi_j d\tau}. \quad (10.11)$$

Since the functions Ψ_j are the exact eigenfunctions of \hat{H} , we can use $\hat{H}\Psi_j = \mathcal{E}_j \Psi_j$ to obtain:

$$E[\Phi] = \frac{\sum_{ij} c_i^* c_j \mathcal{E}_j \int \Psi_i^* \Psi_j d\tau}{\sum_{ij} c_i^* c_j \int \Psi_i^* \Psi_j d\tau}. \quad (10.12)$$

Now using the fact that eigenfunctions of a Hermitian operator form an orthonormal set (or can be made to do so), we can write:

$$E[\Phi] = \frac{\sum_i c_i^* c_i \mathcal{E}_i}{\sum_i c_i^* c_i}. \quad (10.13)$$

We now subtract the exact ground state energy \mathcal{E}_0 from both sides to obtain

$$E[\Phi] - \mathcal{E}_0 = \frac{\sum_i c_i^* c_i (\mathcal{E}_i - \mathcal{E}_0)}{\sum_i c_i^* c_i}. \quad (10.14)$$

Since every term on the right-hand side is greater than or equal to zero, the left-hand side must also be greater than or equal to zero:

$$E[\Phi] \geq \mathcal{E}_0. \quad (10.15)$$

In other words, the energy of any approximate wave function is always greater than or equal to the exact ground state energy \mathcal{E}_0 .

This explains the strategy of the variational method: since the energy of any approximate trial function is always above the true energy, then any variations in the trial function which lower its energy are necessarily making the approximate energy closer to the exact answer. (The trial wave function is also a better approximation to the true ground state wave function as the energy is lowered, although not necessarily in every possible sense unless the limit $\Phi = \Psi_0$ is reached).

10.2.2 Approximated solution for the helium atom

We now have all the ingredients to attempt the simplest approximated solution to the TISEq of a many-electron atom. We can start by writing the total wave function using the Slater determinant of eq. (10.8) in terms of spin-orbitals:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = |\chi_1, \chi_2, \dots, \chi_N\rangle = |\psi_1\phi_\uparrow, \psi_1\phi_\downarrow, \dots, \psi_{\frac{N}{2}}\phi_\uparrow, \psi_{\frac{N}{2}}\phi_\downarrow\rangle, \quad (10.16)$$

and then we can replace it into the TISEq for an N -electron system. This results into a set of N one-electron equations, one for each electron. When we attempt to solve each individual equation, however, we end up with a problem, since the potential energy in the Hamiltonian of eq. (10.1) does not have spherical symmetry because of the electron-electron repulsion term. As such, the one-electron TISEq cannot be simply solved in spherical polar coordinates, as we did for the hydrogen atom in chapter 5. The simplest way of circumventing the problem is to neglect the electron-electron repulsion term (i.e., assume that the electrons are not correlated and do not interact with each other). For a 2-electron atom this procedure is straightforward, since the Hamiltonian can be written as a sum of one-electron Hamiltonians:

$$\hat{H} = \hat{H}_1 + \hat{H}_2, \quad (10.17)$$

with \hat{H}_1 and \hat{H}_2 looking identical to those used in the TISEq of the hydrogen atom. This one-particle Hamiltonian does not depend on the spin of the

electron, and therefore, we can neglect the spin component of the Slater determinant and write the total wave function for the ground state of helium, eq. (10.5), simply as:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2). \quad (10.18)$$

The overall TISEq reduces to a set of two single-particle equations:

$$\begin{aligned}\hat{H}_1\psi_{100}(\mathbf{r}_1) &= E_1\psi_{100}(\mathbf{r}_1) \\ \hat{H}_2\psi_{100}(\mathbf{r}_2) &= E_2\psi_{100}(\mathbf{r}_2),\end{aligned} \quad (10.19)$$

which can then be solved similarly to those for the hydrogen atom, and the solution be combined to give:

$$E = E_1 + E_2. \quad (10.20)$$

In other words, the resulting energy eigenvalue for the ground state of the helium atom in this approximation is equal to twice the energy of a ψ_{100} , $1s$, orbital. The resulting approximated value for the energy of the helium atom is 7,217 kJ/mol, compared with the exact value of 7,620 kJ/mol.

The nuclear charge Z in the ψ_{100} orbital can be used as a variational parameter in the variational method to obtain a more accurate value of the energy. This method provides a result for the ground-state energy of the helium atom of 7,478 kJ/mol (only 142 kJ/mol lower than the exact value), with the nuclear charge parameter minimized at $Z_{\min} = 1.6875$. This new value of the nuclear charge can be interpreted as the effective nuclear charge that is felt by one electron when a second electron is present in the atom. This value is lower than the real nuclear charge ($Z = 2$) because the interaction between the electron and the nuclei is shielded by presence of the second electron.

This procedure can be extended to atoms with more than two electrons, resulting in the so-called Hartree-Fock method. The procedure, however, is not straightforward. We will explain it in more details in the next chapter, since it is the simplest approximation that also describes the chemical bond.

10.3 Chapter Review

10.3.1 Study Questions

1. When going from hydrogen to many-electron atoms, which new term in the electronic hamiltonian prevents an analytic solution of the tise?

- a. electron–nuclear attraction
 - b. kinetic energy
 - c. electron–electron repulsion
 - d. nuclear–nuclear repulsion
 - e. spin–orbit coupling
-

2. For an n -electron atom in non-relativistic approximation, how does the dimensionality of the electronic configuration space scale with n ?

- a. it remains 3 (independent of n)
 - b. it is $3n$ (three spatial coordinates for each electron)
 - c. it is n
 - d. it is $2n$ (one radial and one angular coordinate per electron)
 - e. it is n^2
-

3. Which symmetry requirement must the total electronic wave function for identical electrons satisfy?

- a. symmetry under any electron exchange
 - b. normalization to 1
 - c. invariance under global phase changes only
 - d. periodicity in each coordinate
 - e. antisymmetry under exchange of any two electrons (Pauli principle)
-

4. How is antisymmetry enforced in practice in approximate many-electron wave functions?

- a. by multiplying one-electron functions
 - b. by adding a constant term to the hamiltonian
 - c. by constructing Slater determinants of spin-orbitals
 - d. by ignoring spin altogether
 - e. by symmetrizing products of orbitals
-

5. The variational principle for an approximate normalized trial function Φ states that the energy functional $E[\Phi]$ is?

- a. always less than the exact ground-state energy
 - b. equal to the exact ground-state energy for any trial function
 - c. always greater than or equal to the exact ground-state energy
 - d. independent of the choice of trial function
 - e. meaningful only for excited states
-

6. In the independent-electron approximation used as a starting point for many-electron atoms, each electron is assumed to move in?

- a. in the field of the nucleus only, ignoring other electrons
 - b. in a mean or effective field generated by the nucleus and the average distribution of all other electrons
 - c. in a uniform external electric field
 - d. in a purely harmonic potential
 - e. in free space without any potential
-

7. Why are hydrogen-like orbitals still useful to build approximate many-electron wave functions?

- a. they are exact solutions for all many-electron atoms
 - b. they form a convenient basis set to expand more realistic orbitals
 - c. they automatically include electron correlation
 - d. they depend only on spin variables
 - e. they enforce the pauli principle without determinants
-

8. Which qualitative effect of electron–electron repulsion on the radial distribution of inner versus outer electrons is discussed in many-electron treatments?

- a. outer electrons become more tightly bound than inner ones
- b. inner electrons shield outer ones, so outer electrons feel a smaller effective nuclear charge and are more diffuse

- c. shielding causes inner electrons to move further from the nucleus
 - d. outer electrons experience the full nuclear charge without screening
 - e. both inner and outer electrons collapse into the nucleus
-

9. Which of the following is *not* an effect explicitly included when one uses a simple central-field approximation for many-electron atoms?

- a. average electron–electron repulsion
 - b. screening of nuclear charge
 - c. detailed instantaneous electron–electron correlation (dynamic correlation)
 - d. effective one-electron orbital picture
 - e. approximate shell structure
-

10. In the context of many-electron wave functions, what is meant by “electron correlation” beyond the independent-particle picture?

- a. changes in nuclear positions
 - b. spin conservation
 - c. the normalization condition
 - d. the fact that exact electronic motion is correlated because electrons avoid each other more than in a mean-field description
 - e. inclusion of relativistic mass corrections
-

Answers: 1.c, 2.b, 3.e, 4.c, 5.c, 6.b, 7.b, 8.b, 9.c, 10.d

Chapter 11

Introduction to Molecules

11.1 The Molecular Hamiltonian

For a molecule, we can decompose the Hamiltonian operator as:

$$\hat{H} = \hat{K}_N + \hat{K}_e + \hat{V}_{NN} + \hat{V}_{eN} + \hat{V}_{ee} \quad (11.1)$$

where we have decomposed the kinetic energy operator into nuclear and electronic terms, \hat{K}_N and \hat{K}_e , as well as the potential energy operator into terms representing the interactions between nuclei, \hat{V}_{NN} , between electrons, \hat{V}_{ee} , and between electrons and nuclei, \hat{V}_{eN} . Each term can then be calculated using:

$$\begin{aligned}
 \hat{K}_N &= - \sum_i^{\text{nuclei}} \frac{\hbar^2}{2M_i} \nabla_{\mathbf{R}_i}^2 \\
 \hat{K}_e &= - \sum_i^{\text{electrons}} \frac{\hbar^2}{2m_e} \nabla_{\mathbf{r}_i}^2 \\
 \hat{V}_{NN} &= \sum_i \sum_{j>i} \frac{Z_i Z_j e^2}{4\pi\epsilon_0 |\mathbf{R}_i - \mathbf{R}_j|} \\
 \hat{V}_{eN} &= - \sum_i \sum_j \frac{Z_i e^2}{4\pi\epsilon_0 |\mathbf{R}_i - \mathbf{r}_j|} \\
 \hat{V}_{ee} &= \sum_i \sum_{i<j} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|},
 \end{aligned} \tag{11.2}$$

where M_i , Z_i , and \mathbf{R}_i are the mass, atomic number, and coordinates of nucleus i , respectively, and all other symbols are the same as those used in eq. (10.1) for the many-electron atom Hamiltonian.

11.1.1 Small terms in the molecular Hamiltonian

The operator in eq. (11.1) is known as the “exact” nonrelativistic Hamiltonian in field-free space. However, it is important to remember that it neglects at least two effects. Firstly, although the speed of an electron in a hydrogen atom is less than 1% of the speed of light, relativistic mass corrections can become appreciable for the inner electrons of heavier atoms. Secondly, we have neglected the spin-orbit effects, which is explained as follows. From the point of view of an electron, it is being orbited by a nucleus which produces a magnetic field (proportional to \mathbf{L}); this field interacts with the electron’s magnetic moment (proportional to \mathbf{S}), giving rise to a spin-orbit interaction (proportional to $\mathbf{L} \cdot \mathbf{S}$ for a diatomic.) Although spin-orbit effects can be important, they are generally neglected in quantum chemical calculations, and we will neglect them in the remainder of this textbook as well.

11.2 The Born-Oppenheimer Approximation

As we already saw in the previous chapter, if a Hamiltonian is separable into two or more terms, then the total eigenfunctions are products of the

individual eigenfunctions of the separated Hamiltonian terms. The total eigenvalues are then sums of individual eigenvalues of the separated Hamiltonian terms.

For example, let's consider a Hamiltonian that is separable into two terms, one involving coordinate q_1 and the other involving coordinate q_2 :

$$\hat{H} = \hat{H}_1(q_1) + \hat{H}_2(q_2) \quad (11.3)$$

with the overall Schrödinger equation being:

$$\hat{H}\psi(q_1, q_2) = E\psi(q_1, q_2). \quad (11.4)$$

If we assume that the total wave function can be written in the form:

$$\psi(q_1, q_2) = \psi_1(q_1)\psi_2(q_2), \quad (11.5)$$

where $\psi_1(q_1)$ and $\psi_2(q_2)$ are eigenfunctions of \hat{H}_1 and \hat{H}_2 with eigenvalues E_1 and E_2 , then:

$$\begin{aligned} \hat{H}\psi(q_1, q_2) &= (\hat{H}_1 + \hat{H}_2)\psi_1(q_1)\psi_2(q_2) \\ &= \hat{H}_1\psi_1(q_1)\psi_2(q_2) + \hat{H}_2\psi_1(q_1)\psi_2(q_2) \\ &= E_1\psi_1(q_1)\psi_2(q_2) + E_2\psi_1(q_1)\psi_2(q_2) \\ &= (E_1 + E_2)\psi_1(q_1)\psi_2(q_2) \\ &= E\psi(q_1, q_2) \end{aligned} \quad (11.6)$$

Thus the eigenfunctions of \hat{H} are products of the eigenfunctions of \hat{H}_1 and \hat{H}_2 , and the eigenvalues are the sums of eigenvalues of \hat{H}_1 and \hat{H}_2 .

If we examine the nonrelativistic Hamiltonian in eq. (11.1), we see that the \hat{V}_{eN} terms prevents us from cleanly separating the electronic and nuclear coordinates and writing the total wave function. If we neglect these terms, we can write the total wave function as:

$$\psi(\mathbf{r}, \mathbf{R}) = \psi_e(\mathbf{r})\psi_N(\mathbf{R}), \quad (11.7)$$

This approximation is called the **Born-Oppenheimer approximation**, and allows us to treat the nuclei as nearly fixed with respect to electron motion. The Born-Oppenheimer approximation is almost always quantitatively

correct, since the nuclei are much heavier than the electrons and the (fast) motion of the latter does not affect the (slow) motion of the former. Using this approximation, we can fix the nuclear configuration at some value, $\mathbf{R}_\mathbf{a}$, and solve for the electronic portion of the wave function, which is dependent only parametrically on \mathbf{R} (we write this wave function as $\psi_e(\mathbf{r}; \mathbf{R}_\mathbf{a})$, where the semicolon indicate the parametric dependence on the nuclear configuration). To solve the TISEq we can then write the electronic Hamiltonian as:

$$\hat{H}_e = \hat{K}_e(\mathbf{r}) + \hat{V}_{eN}(\mathbf{r}; \mathbf{R}_\mathbf{a}) + \hat{V}_{ee}(\mathbf{r}) \quad (11.8)$$

where we have also factored out the nuclear kinetic energy, \hat{K}_N (since it is smaller than \hat{K}_e by a factor of $\frac{M_i}{m_e}$), as well as $\hat{V}_{NN}(\mathbf{R})$. This latter approximation is justified, since in the Born-Oppenheimer approximation \mathbf{R} is just a parameter, and $\hat{V}_{NN}(\mathbf{R}_\mathbf{a})$ is a constant that shifts the eigenvalues only by some fixed amount. This electronic Hamiltonian results in the following TISEq:

$$\hat{H}_e \psi_e(\mathbf{r}; \mathbf{R}_\mathbf{a}) = E_e \psi_e(\mathbf{r}; \mathbf{R}_\mathbf{a}), \quad (11.9)$$

which is the equation that is used to explain the chemical bond in the next section. Notice that eq. (11.9) is not the total TISEq of the system, since the nuclear eigenfunction and its eigenvalues (which can be obtained solving the Schrödinger equation with the nuclear Hamiltonian) are neglected. As a final note, in the remainder of this textbook we will confuse the term “total energy” with “total energy at fixed geometry”, as is customary in many other quantum chemistry textbooks (i.e., we are neglecting the nuclear kinetic energy). This is just E_e of eq. (11.9), plus the constant shift, $\hat{V}_{NN}(\mathbf{R}_\mathbf{a})$, given by the nuclear-nuclear repulsion.

11.3 Solving the Electronic Eigenvalue Problem

Once we have invoked the Born-Oppenheimer approximation, we can attempt to solve the electronic TISEq in eq. (11.9). However, for molecules with more than one electron, we need to—once again—keep in mind the antisymmetry of the wave function. This obviously means that we need to write the electronic wave function as a Slater determinant (i.e., all molecules but H_2^+ and a few related highly exotic ions). Once this is done, we can work

on approximating the Hamiltonian, a task that is necessary because the presence of the electron-electron repulsion term forbids its analytic treatment. Similarly to the many-electron atom case, the simplest approximation to solve the molecular electronic TISEq is to use the variational method and to neglect the electron-electron repulsion. As we noticed in the previous chapter, this approximation is called the Hartree-Fock method.

11.3.1 The Hartree-Fock Method

The main difference when we apply the variational principle to a molecular Slater determinant is that we need to build orbitals (one-electron wave functions) that encompass the entire molecule. This can be done by assuming that the atomic contributions to the molecular orbitals will closely resemble the orbitals that we obtained for the hydrogen atom. The total molecular orbital can then be built by linearly combine these atomic contributions. This method is called **linear combination of atomic orbitals (LCAO)**. A consequence of the LCAO method is that the atomic orbitals on two different atomic centers are not necessarily orthogonal, and eq. (10.12) cannot be simplified easily. If we replace each atomic orbital $\psi(\mathbf{r})$ with a linear combination of suitable basis functions $f_i(\mathbf{r})$:

$$\psi(\mathbf{r}) = \sum_i^m c_i f_i(\mathbf{r}), \quad (11.10)$$

we can then use the following notation:

$$H_{ij} = \int \phi_i^* \hat{H} \phi_j d\tau, \quad S_{ij} = \int \phi_i^* \phi_j d\tau, \quad (11.11)$$

to simplify eq. (10.12) to:

$$E[\Phi] = \frac{\sum_{ij} c_i^* c_j H_{ij}}{\sum_{ij} c_i^* c_j S_{ij}}. \quad (11.12)$$

Differentiating this energy with respect to the expansion coefficients c_i yields a non-trivial solution only if the following “secular determinant” equals zero:

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} & \cdots & H_{1m} - ES_{1m} \\ H_{21} - ES_{21} & H_{22} - ES_{22} & \cdots & H_{2m} - ES_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ H_{m1} - ES_{m1} & H_{m2} - ES_{m2} & \cdots & H_{mm} - ES_{mm} \end{vmatrix} = 0 \quad (11.13)$$

where m is the number of basis functions used to expand the atomic orbitals. Solving this set of equations with a Hamiltonian where the electron-electron correlation is neglected results in non-trivial, but possible. The reason for the complications comes from the fact that even if we are neglecting the direct interaction between electrons, each of them interact with the nuclei through an interaction that is screened by the average field of all other electrons, similarly to what we saw for the helium atom. This means that the Hamiltonian itself and the value of the coefficients c_i in the wave function mutually depend on each other. A solution to this problem can be achieved numerically using specialized computer programs that use a cycle called the **self-consistent-field (SCF) procedure**. Starting from an initial guess of the coefficients, an approximated Hamiltonian operator is built from them and used to solve eq. (11.13). This solution gives updated values of the coefficients, which can then be used to create an improved version of the approximated Hamiltonian. This procedure is repeated until both the coefficients and the operator do not change anymore. From this final solution, the energy of the molecule is then calculated.

11.4 Chapter Review

11.4.1 Study Questions

1. In the full non-relativistic molecular Hamiltonian (before approximations), which terms are present?
 - a. only electronic kinetic energy and electron–nuclear attraction
 - b. only nuclear kinetic energy and nuclear–nuclear repulsion
 - c. spin–orbit and relativistic corrections only

- d. only electron–electron repulsion and nuclear–nuclear repulsion
 - e. electronic and nuclear kinetic energies, electron–nuclear, electron–electron, and nuclear–nuclear interactions
-

2. What makes the exact molecular Schrödinger equation harder than the atomic case?

- a. electrons disappear in molecules
 - b. nuclei can be treated classically in atoms but not in molecules
 - c. the presence of multiple moving nuclei in addition to many electrons
 - d. there is no coulomb interaction in molecules
 - e. the electron mass changes in molecules
-

3. What is the central physical idea behind the Born–Oppenheimer approximation?

- a. electrons are much heavier than nuclei
 - b. nuclei are much heavier than electrons, so electrons adjust almost instantaneously to nuclear motion
 - c. electrons and nuclei move with the same characteristic time scale
 - d. both electrons and nuclei can be treated as classical particles
 - e. nuclei do not contribute to the energy at all
-

4. Mathematically, the Born–Oppenheimer approximation relies on what property of the Hamiltonian?

- a. exact commutation with the position operator
 - b. separability into electronic and nuclear parts (plus interactions)
 - c. absence of kinetic energy terms
 - d. Hermiticity of the potential energy operator
 - e. linearity of the time derivative
-

5. In practical quantum-chemistry, what is usually meant (by abuse of language) by “total energy” of a molecule at fixed geometry?

- a. electronic energy E_e at that geometry plus the constant nuclear–nuclear repulsion term
 - b. nuclear kinetic energy only
 - c. electronic energy without nuclear–nuclear repulsion
 - d. purely classical potential energy
 - e. nuclear vibrational zero-point energy
-

6. Which of the following statements regarding potential-energy surfaces (PESs) is correct?

- a. a PES is the graph of electronic energy as a function of nuclear geometry
- b. a PES is the graph of nuclear kinetic energy versus time
- c. a PES is independent of nuclear positions
- d. a PES depends only on electron spin

- e. PESs are defined only for diatomic molecules
-

7. Which of the following is *not* a direct consequence or assumption of the Born–Oppenheimer approximation?

- a. electronic and nuclear motions can be approximately separated
 - b. electrons move in the field of fixed nuclei
 - c. nuclear motion can be described on a pes obtained from electronic energies
 - d. electron–nuclear coupling can never be important for any property
 - e. total wavefunctions are approximated as products of electronic and nuclear parts
-

8. When the variational principle is applied to a Slater determinant built from molecular spin-orbitals, minimizing the energy with respect to the orbital coefficients leads to what?

- a. the time-dependent Schrödinger equation
 - b. an over-determined system with no nontrivial solutions
 - c. a single algebraic equation for the total energy
 - d. purely classical equations of motion for nuclei
 - e. a set of coupled one-electron equations involving the fock operator (Hartree–Fock equations)
-

9. Which of the following sentences describes the Fock operator \hat{F} in Hartree–Fock theory?

- a. a two-electron operator containing only electron–electron repulsion
 - b. a purely nuclear operator
 - c. a one-electron effective operator containing kinetic, nuclear attraction, and averaged electron–electron (Coulomb and exchange) terms
 - d. an operator that depends only on spin and not on spatial coordinates
 - e. the exact electronic Hamiltonian
-

10. In the Hartree–Fock method, how is the Pauli exclusion principle enforced for many-electron systems?

- a. by forcing all electrons into different nuclei
 - b. by requiring all spatial orbitals to be orthogonal but ignoring spin
 - c. by constructing an antisymmetric Slater determinant of spin-orbitals
 - d. by adding a large repulsive potential at short inter-electronic distances
 - e. by using only hydrogen-like orbitals
-

Answers: 1.e, 2.c, 3.b, 4.b, 5.a, 6.a, 7.d, 8.e, 9.c, 10.c

Chapter 12

The Chemical Bond in Diatomic Molecules

In this chapter we will see a couple of examples of how the concept and mathematics of quantum mechanics can be applied to understand the chemical bond in molecules. We will start from the simplest molecule, the H_2^+ molecular ion, and then we will move on to the simplest two-electron bond in the hydrogen molecule. To simplify the notation in this chapter, we will move away from S.I. units and use a set tailored for molecules, called *atomic units* (*a.u.*). This set of units is built by setting $\hbar = e = m_e = a_0 = 1$. As an example of the simplification that *a.u.* allows, the energy eigenvalues of the hydrogen atom, eq. (5.8), simply becomes $E_n = -\frac{1}{2n^2}$ in the *a.u.* of energy, which are called Hartrees.

12.1 The Chemical Bond in the Hydrogen Molecular Cation

This system has only one electron, but since its geometry is not spherical (figure 12.1), the TISEq cannot be solved analytically as for the hydrogen atom.

The electron is at point P , while the two protons are at position A and B at a fixed distance R . Using the Born-Oppenheimer approximation we can write the one-electron molecular Hamiltonian in *a.u.* as:

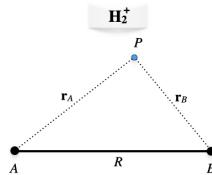


Figure 12.1: Geometry of the hydrogen molecular cation.

$$\hat{H} = \hat{H}_e + \frac{1}{R} = \left(-\frac{1}{2} \nabla^2 - \frac{1}{\mathbf{r}_A} - \frac{1}{\mathbf{r}_B} \right) + \frac{1}{R} \quad (12.1)$$

As a first approximation to the variational wave function, we can build the one-electron molecular orbital (MO) by linearly combine two $1s$ hydrogenic orbitals centered at A and B , respectively:

$$\varphi = c_1 a + c_2 b, \quad (12.2)$$

with:

$$\begin{aligned} a &= 1s_A = (\psi_{100})_A \\ b &= 1s_B = (\psi_{100})_B. \end{aligned} \quad (12.3)$$

Using eq. (11.11) and considering that the nuclei are identical, we can define the integrals $H_{aa} = H_{bb}$, $H_{ab} = H_{ba}$ and $S_{ab} = S$ (while $S_{aa} = 1$ because the hydrogen atom orbitals are normalized). The secular equation, eq. (11.13) can then be written:

$$\begin{vmatrix} H_{aa} - E & H_{ab} - ES \\ H_{ab} - ES & H_{aa} - E \end{vmatrix} = 0 \quad (12.4)$$

\end{equation}

The expansion of the determinant results into:

$$(H_{aa} - E)^2 = (H_{ab} - ES)^2 \\ H_{aa} - E = \pm(H_{ab} - ES), \quad (12.5)$$

with roots:

$$E_+ = \frac{H_{aa} + H_{ab}}{1 + S} = H_{aa} + \frac{H_{ba} - SH_{aa}}{1 + S}, \\ E_- = \frac{H_{aa} - H_{ab}}{1 - S} = H_{aa} - \frac{H_{ba} - SH_{aa}}{1 - S}, \quad (12.6)$$

the first corresponding to the ground state, the second to the first excited state. Solving for the best value for the coefficients of the linear combination for the ground state E_+ , we obtain:

$$c_1 = c_2 = \frac{1}{\sqrt{2 + 2S}}, \quad (12.7)$$

which gives the bonding MO:

$$\varphi_+ = \frac{a + b}{\sqrt{2 + 2S}}. \quad (12.8)$$

Proceeding similarly for the excited state, we obtain:

$$c_1 = \frac{1}{\sqrt{2 - 2S}} \quad c_2 = -\frac{1}{\sqrt{2 - 2S}}, \quad (12.9)$$

which gives the antibonding MO:

$$\varphi_- = \frac{b - a}{\sqrt{2 - 2S}}. \quad (12.10)$$

These results can be summarized in the molecular orbital diagram of figure 12.2 We notice that the splitting of the doubly degenerate atomic level under the interaction is non-symmetric for $S \neq 0$, the antibonding level being more repulsive and the bonding less attractive than the symmetric case occurring for $S = 0$.

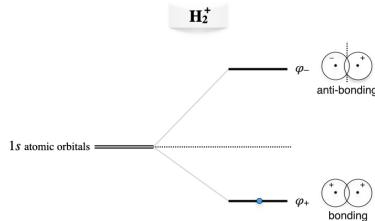


Figure 12.2: Molecular orbitals diagram for the hydrogen molecular cation.

Calculating the values for the integrals and repeating these calculations for different internuclear distances, R , results in the plot of figure 12.3 As we see from the plots, the ground state solution is negative for a vast portion of the plot. The energy is negative because the electronic energy calculated with the bonding orbital is lower than the nuclear repulsion. In other words, the creation of the molecular orbital stabilizes the molecular configuration versus the isolated fragments (one hydrogen atom and one proton).

12.2 The Chemical Bond in the Hydrogen Molecule

We can now examine the formation of the two-electron chemical bond in the H_2 molecule. With reference to figure 12.4, the molecular Hamiltonian for H_2 in a.u. in the Born-Oppenheimer approximation will be:

$$\begin{aligned}
 \hat{H} &= \hat{H}_e + \frac{1}{R} \\
 &= \left(-\frac{1}{2} \nabla_1^2 - \frac{1}{\mathbf{r}_{A1}} - \frac{1}{\mathbf{r}_{B1}} \right) + \left(-\frac{1}{2} \nabla_2^2 - \frac{1}{\mathbf{r}_{A2}} - \frac{1}{\mathbf{r}_{B2}} \right) + \frac{1}{r_{12}} + \frac{1}{R} \\
 &= \hat{h}_1 + \hat{h}_2 + \frac{1}{r_{12}} + \frac{1}{R},
 \end{aligned} \tag{12.11}$$

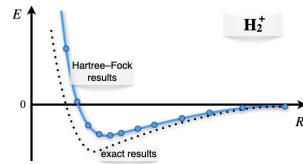


Figure 12.3: Born-Oppenheimer energy landscape for the hydrogen molecular cation.

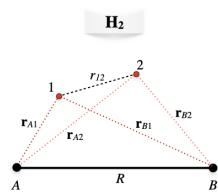


Figure 12.4: Geometry of the hydrogen molecule.

where \hat{h} is the one-electron Hamiltonian. As for the previous case, we can build the first approximation to the molecular wave function by considering two $1s$ atomic orbitals $a(\mathbf{r}_1)$ and $b(\mathbf{r}_2)$ centered at A and B , respectively, having an overlap S . If we Neglect the electron-electron repulsion term, $\frac{1}{r_{12}}$, the resulting Hartree-Fock equations are exactly the same as in the previous case. The most important difference, though, is that in this case we need to consider the spin of the two electrons. Proceeding similarly to what we have done for the many-electron atom in chapter 10, we can build an antisymmetric wave function for H_2 using a Slater determinant of doubly occupied MOs. For the ground state, we can use the lowest energy orbital obtained from the solution of the Hartree-Fock equations, which we already obtained in eq. (12.8). Using a notation that is based on the symmetry of the molecule, this bonding orbital in H_2 is usually called σ_g , where σ refers to the σ bond that forms between the two atoms. The Slater determinant for the ground state is therefore:¹

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = |\sigma_g \phi_\uparrow, \sigma_g \phi_\downarrow\rangle, = \sigma_g(\mathbf{r}_1)\sigma_g(\mathbf{r}_2) \frac{1}{\sqrt{2}} [\phi_\uparrow \phi_\downarrow - \phi_\downarrow \phi_\uparrow], \quad (12.12)$$

where:

$$\sigma_g = \varphi_+ = \frac{(\psi_{100})_A + (\psi_{100})_B}{\sqrt{2 + 2S}}. \quad (12.13)$$

The energies and the resulting MO diagram is similar to that for H_2^+ , with the only difference that two electron will be described by the same σ_g MO (figure 12.5).

As for the many-electron atoms, the Hartree-Fock method is just an approximation to the exact solution. The accurate theoretical value for the bond energy at the bond distance of $R_e = 1.4 a_0$ is $E = -0.17447 E_h$. The variational result obtained with the wave function in eq. (12.12) is $E = -0.12778 E_h$, which is $\sim 73\%$ of the exact value. The variational coefficient (i.e., the orbital exponent, c_0 , that enters the $1s$ orbital formula $\psi_{100} = \frac{1}{\pi} \exp[c_0 r]$) is optimized at $c_0 = 1.1695$, a value that shows how the orbitals significantly contract due to spherical polarization.

If we scan the Born-Oppenheimer energy landscape using the wave function in eq. (12.12) as we have done for H_2^+ , we obtain the plot in figure 12.6.

¹Compare this equation to (10.6) for the helium atom.

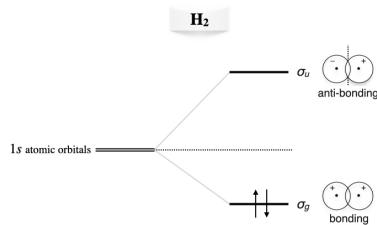


Figure 12.5: Molecular orbitals diagram for the hydrogen molecule.

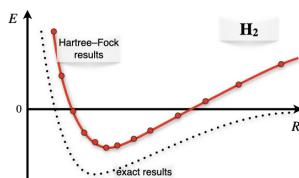


Figure 12.6: Born-Oppenheimer energy landscape for the hydrogen molecule.

As we can see, the Hartree-Fock results for H₂ describes the formation of the bond qualitatively around the bond distance (minimum of the curve), but they fail to describe the molecule at dissociation. This happens because in eq. (12.12) both electrons are in the same orbital with opposite spin (electrons are coupled), and the orbital is shared among both centers. At dissociation, this corresponds to an erroneous ionic dissociation state where both electron are localized on either one of the two centers (this center is therefore negatively charged), with the other proton left without electrons. This is in contrast with the correct dissociation, where each electron should be localized around each center (and therefore, it should be uncoupled from the other electron). This error is once again the result of the approximations that are necessary to treat the TISEq of a many-electron system. It is obviously not a failure of quantum mechanics, and it can be easily corrected using more accurate approximations on modern computers.

12.3 Chapter Review

12.3.1 Study Questions

1. In the H₂⁺ treatment, which one-electron molecular Hamiltonian (in atomic units, and using the Born–Oppenheimer approximation) is used?

- a. $\hat{H} = -\frac{1}{2}\nabla^2 - \frac{1}{r_a}$
 - b. $\hat{H} = -\frac{1}{2}\nabla^2 - \frac{1}{r_b}$
 - c. $\hat{H} = -\frac{1}{2}\nabla^2 - \frac{1}{r_a} - \frac{1}{r_b}$
 - d. $\hat{H} = -\frac{1}{2}\nabla^2 + \frac{1}{r_a} + \frac{1}{r_b}$
 - e. $\hat{H} = -\nabla^2 + \frac{1}{r_a r_b}$
-

2. Which of the following is used as the first variational ansatz for the one-electron molecular orbital in H₂⁺?

- a. $\varphi = a$
 - b. $\varphi = b$
 - c. $\varphi = ab$
 - d. $\varphi = c_1a + c_2b$
 - e. $\varphi = a^2 + b^2$
-

3. In the secular problem for H_2^+ , what is the physical meaning of the overlap integral $S = S_{ab}$?

- a. integral $\int a(\mathbf{r})b(\mathbf{r})d\mathbf{r}$ measuring non-orthogonality of the two 1s orbitals
 - b. nuclear repulsion between the two protons
 - c. kinetic energy between a and b
 - d. electron-electron repulsion
 - e. spin overlap between alpha and beta functions
-

4. The diagonalization of the 2×2 secular problem for H_2^+ gives which types of molecular orbitals?

- a. a nonbonding and an antinonbonding orbital
 - b. a bonding φ_+ and an antibonding φ_- combination
 - c. two identical orbitals with the same energy
 - d. two core orbitals localized on each nucleus
 - e. two purely antibonding orbitals
-

5. What is the physical meaning of the negative total energy of the ground state for a range of r in H_2^+ ?

- a. the molecular ion is unstable relative to separated $\text{H} + \text{H}^+$
 - b. the molecular orbital stabilizes the system relative to the separated fragments
 - c. electron–nuclear attraction is negligible
 - d. only nuclear–nuclear repulsion matters
 - e. there is no chemical bond
-

6. Which of the following is the correct molecular Hamiltonian, in a.u. and under the Born–Oppenheimer approximation, for H_2 ?

- a. $\hat{h}_1 + \hat{h}_2$ only
 - b. $\hat{h}_1 + \hat{h}_2 - 1/r_{12}$
 - c. $\hat{k}_n + \hat{k}_e$ only
 - d. $1/r_{12} + 1/r$ only
 - e. $\hat{h}_1 + \hat{h}_2 + 1/r_{12} + 1/r$
-

7. For the ground state of H_2 in the minimal MO picture used in this chapter, which orbital and occupancy pattern are used?

- a. two electrons singly occupying φ_+ and φ_- with parallel spins
- b. both electrons in the antibonding MO φ_- with opposite spin
- c. one electron in σ_g , one in a core orbital

d. both electrons in the bonding MO $\sigma_g = \varphi_+$ with opposite spin

e. both electrons in a localized atomic orbital on a single nucleus

8. why does the RHF-type description of H₂ fail at large H–H separations (dissociation limit)?

a. it breaks spin symmetry

b. it forces both electrons to remain in the same delocalized orbital shared between centers, rather than localizing one electron on each atom

c. it neglects nuclear repulsion

d. it neglects electron kinetic energy

e. it overestimates overlap as $S = 1$

9. What is the main qualitative effect of including electron–electron repulsion and better correlation methods beyond the simple Hartree–Fock picture in H₂?

a. they destroy the bond minimum

b. they force both electrons into antibonding orbitals

c. they improve the description of the potential-energy curve, particularly the dissociation region

d. they eliminate spin considerations

e. they make the electrons behave classically

10. Comparing H_2^+ and H_2 , which statement about their bonding descriptions is true?

- a. both are one-electron systems with identical Hamiltonians
- b. H_2^+ is exactly solvable analytically while H_2 is not
- c. H_2^+ has one-electron bonding described by a single MO; H_2 requires a two-electron antisymmetric wave function including spin
- d. H_2^+ has no bound state
- e. H_2 can be described without consideration of spin

Answers: 1.c, 2.d, 3.a, 4.b, 5.b, 6.e, 7.d, 8.b, 9.c, 10.c

Chapter 13

The Chemical Bond in Polyatomic Molecules

The structure in space of polyatomic molecules depends on the stereochemistry of their chemical bonds and can be determined by solving the (approximated) TISEq using the Born—Oppenheimer approximation using a method that uses a linear combination of atomic orbitals to form molecular orbitals (LCAO-MO).

13.1 The Chemical Bond in the Water Molecule Using a Minimal Basis

For a minimal representation of the two hydrogen atoms, we need two $1s$ functions, one centered on each atom. Oxygen has electrons in the second principal quantum level, so we will need one $1s$, one $2s$, and three $2p$ functions (one each of p_x , p_y , and p_z). Summarizing, for a minimal representation of the water wave function we need five orbitals on oxygen, plus one each on the hydrogen atoms, for a total of 7 functions. From these atomic functions, we can build a total wave function using the LCAO method of chapter 11, and then we can use the variational principle, in conjunction with the Hartree—Fock (HF) method, to build and solve a secular determinant that looks similar to that in eq. (11.13), with $m = 7$ being the total number of basis functions. The approximated Hamiltonian operator in the HF method is called the Fock operator, and it can be divided into

one-electron integrals, comprising the kinetic and potential energy contributions:

$$\begin{aligned} K_{ij} &= \int \phi_i^* \hat{K} \phi_j \, d\tau = \int \phi_i^* \left(-\frac{1}{2} \nabla^2 \right) \phi_j \, d\tau \\ V_{ij} &= \int \phi_i^* \hat{V} \phi_j \, d\tau = \int \phi_i^* \left(- \sum_k^{\text{nuclei}} \frac{Z_k}{r_k} \right) \phi_j \, d\tau, \end{aligned} \quad (13.1)$$

as well as two-electron integrals describing the coulomb repulsion between electrons:

$$V_{ijkl} = \iint \phi_i^* \phi_j^* \hat{r}_{12} \phi_k \phi_l \, d\tau_1 d\tau_2 = \iint \phi_i^* \phi_j^* \left(\frac{1}{r_{12}} \right) \phi_k \phi_l \, d\tau_1 d\tau_2. \quad (13.2)$$

Despite the minimal basis set, the total number of integrals that need to be calculated for water is large, since i , j , k , and l can be any one of the 7 basis functions. Hence there are $7 \times 7 = 49$ kinetic energy integrals, and the same number of potential energy integrals for each nucleus, resulting in $7 \times 7 \times 3 = 147$. The grand total of one-electron integrals is thus 196. For the two-electron integrals, we have $7 \times 7 \times 7 \times 7 = 2,401$ integrals to calculate. Overall for this simple calculation on water, we need almost 2,600 integrals.¹

All this to find 5 occupied molecular orbitals from which to form a final Slater determinant (10 electrons, two to an orbital, so 5 orbitals). The situation sounds horrible, but it should be recognized that the solutions to all of the integrals are known to be analytic formulae involving only interatomic distances, cartesian exponents, and the values of a single exponent in the

¹The numbers computed here involve the minimum amount of uncontracted “hydrogenic” functions that can be used for calculation on water. In real-life calculations a linear combination of simpler primitive functions (gaussians) is used to describe a single uncontracted function. For example in the simplest case, the STO-3G basis set, each uncontracted function is composed of 3 primitive gaussian functions. Thus, for any individual one-electron integral, there will be $3 \times 3 = 9$ separate integrals involving the primitives. There are thus $9 \times 196 = 1,764$ individual primitive one-electron integrals. As for the two-electron integrals, again, every individual integral will require considering every possible combination of constituent primitives which is $3 \times 3 \times 3 \times 3 = 81$. Thus, the total number of primitive two-electron integrals is $81 \times 2,401 = 194,481$ (gulp!). Notice that even for this small molecule the number of two-electron integrals totally dominates the number of one-electron integrals. The disparity only increases with molecular size. Notice: Portions of this section are based on Prof. C.J. Cramer's lecture notes available (here)[\[http://pollux.chem.umn.edu/4502/3502_lecture_29.pdf\]](http://pollux.chem.umn.edu/4502/3502_lecture_29.pdf)

atomic functions. If we use slightly simpler gaussian functions instead of the more complicated hydrogenic solutions, the total number of floating-point operations to solve the integrals is roughly 1,000,000. In computer speak that's one megaflop (megaflop = million FLoating-point OPerations). A modern digital computer processor can achieve gigaflop per second performance, so the computer can accomplish all these calculations in under one second. An additional way in which things can be improved is to recognize that the molecule has symmetries that can be exploited to reduce the number of total integrals that needs to be calculated.

13.2 Hartree-Fock Calculation for Water

To find the Hartree-Fock (HF) molecular orbitals (MOs) we need to solve the following secular determinant:

$$\begin{vmatrix} F_{11} - ES_{11} & F_{12} - ES_{12} & \cdots & F_{17} - ES_{17} \\ F_{21} - ES_{21} & F_{22} - ES_{22} & \cdots & F_{27} - ES_{27} \\ \vdots & \vdots & \ddots & \vdots \\ F_{71} - ES_{71} & F_{72} - ES_{72} & \cdots & F_{77} - ES_{77} \end{vmatrix} = 0 \quad (13.3)$$

with S_{ij} being the overlap integrals of eq. (11.11), and F_{ij} the matrix elements of the Fock operator, defined using the one- and two-electron integrals in eq. (13.1) and eq. (13.2) as:

$$F_{ij} = K_{ij} + V_{ij} + \sum_{kl} P_{kl} \left[V_{ijkl} - \frac{1}{2} V_{ikjl} \right], \quad (13.4)$$

with the density matrix elements P_{kl} defined as:

$$P_{kl} = 2 \sum_i^{\text{occupied}} a_{ki} a_{li}, \quad (13.5)$$

where the a values are the coefficients of the basis functions in the occupied molecular orbitals. These values will be determined using the SCF procedure, which proceeds as follows: At the first step we simply guess what these

are, then we iterate through solution of the secular determinant to derive new coefficients and we continue to do so until self-consistency is reached (i.e. the $N+1$ step provides coefficients and energies that are equal to those in the N step).

We can try to solve the SCF procedure for water using a fixed geometry of the nuclei close to the experimental structure: O-H bond lengths of 0.95 \AA and a valence bond angle at oxygen of 104.5° . To do so, we can use a minimal basis functions composed of the following seven orbitals: basis function #1 is an oxygen $1s$ orbital, #2 is an oxygen $2s$ orbital, #3 is an oxygen $2p_x$ orbital, #4 is an oxygen $2p_y$ orbital, #5 is an oxygen $2p_z$ orbital, #6 is one hydrogen $1s$ orbital, and #7 is the other hydrogen $1s$ orbital. The corresponding integrals introduced in the previous section can be calculated using a quantum chemistry code. The calculated overlap matrix elements are:

$$\mathbf{S} = \begin{bmatrix} \text{O } 1s & \text{O } 2s & \text{O } 2p_x & \text{O } 2p_y & \text{O } 2p_z & \text{H}_a \text{ } 1s & \text{H}_b \text{ } 1s \\ 1.000 & & & & & & \\ 0.237 & 1.000 & & & & & \\ 0.000 & 0.000 & 1.000 & & & & \\ 0.000 & 0.000 & 0.000 & 1.000 & & & \\ 0.000 & 0.000 & 0.000 & 0.000 & 1.000 & & \\ 0.055 & 0.479 & 0.000 & 0.313 & -0.242 & 1.000 & \\ 0.055 & 0.479 & 0.000 & -0.313 & -0.242 & 0.256 & 1.000 \end{bmatrix} \begin{array}{l} \text{O } 1s \\ \text{O } 2s \\ \text{O } 2p_x \\ \text{O } 2p_y \\ \text{O } 2p_z \\ \text{H}_a \text{ } 1s \\ \text{H}_b \text{ } 1s \end{array} \quad (13.6)$$

There are many noteworthy features in \mathbf{S} . First, it is shown in a lower packed triangular form because every element j, i is the same as the element i, j by symmetry, and every diagonal element is 1 because the basis functions are normalized. Note that, again by symmetry, every p orbital on oxygen is orthogonal (overlap = zero) with every s orbital and with each other, but the two s orbitals do overlap (this is due to the fact that they are not pure hydrogenic orbitals—which would indeed be orthogonal—but they have been optimized, so $S_{12} = 0.237$). Note also that the oxygen $1s$ orbital overlaps about an order of magnitude less with any hydrogen $1s$

orbital than does the oxygen 2s orbital, reflecting how much more rapidly the first quantum-level orbital decays compared to the second. Note that by symmetry the oxygen p_x cannot overlap with the hydrogen 1s functions (positive overlap below the plane exactly cancels negative overlap above the plane) and that the oxygen p_y overlaps with the two hydrogen 1s orbitals equally in magnitude but with different sign because the p orbital has different phase at its different ends. Finally, the overlap of the p_z is identical with each H 1s because it is not changing which lobe it uses to interact. The kinetic energy matrix (in a.u.) is:

$$\mathbf{K} = \begin{bmatrix} 29.003 & & & & & & \\ -0.168 & 0.808 & & & & & \\ 0.000 & 0.000 & 2.529 & & & & \\ 0.000 & 0.000 & 0.000 & 2.529 & & & \\ 0.000 & 0.000 & 0.000 & 0.000 & 2.529 & & \\ -0.002 & 0.132 & 0.000 & 0.229 & -0.177 & 0.760 & \\ -0.002 & 0.132 & 0.000 & -0.229 & -0.177 & 0.009 & 0.760 \end{bmatrix} \quad (13.7)$$

Notice that every diagonal term is much larger than any off-diagonal term. Recall that each kinetic energy integral, eq. (13.1), involves the Laplacian operator, ∇^2 . The Laplacian reports back the sum of second derivatives in all coordinate directions. That is, it is a measure of how fast the slope of the function is changing in various directions. If we take two atomic orbitals μ and ν far apart from each other, then since gaussians go to zero at least exponentially fast with distance, ν is likely to be very flat where μ is large. The second derivative of a flat function is zero. So, every point in the integration will be roughly the amplitude of μ times zero, and not much will accumulate. For the diagonal element, on the other hand, the interesting second derivatives will occur where the function has maximum amplitude (amongst other places) so the accumulation should be much larger. Notice also that off-diagonal terms can be negative. That is because there is no real physical meaning to a kinetic energy expectation value involving two different orbitals. It is just an integral that appears in the complete secular determinant. Symmetry again keeps p orbitals from mixing with s orbitals or with each other. The nuclear attraction matrix is:

$$\mathbf{V} = \begin{bmatrix} -61.733 & & & & & & \\ -7.447 & -10.151 & & & & & \\ 0.000 & 0.000 & -9.926 & & & & \\ 0.000 & 0.000 & 0.000 & -10.152 & & & \\ 0.000 & 0.000 & 0.000 & 0.000 & -10.088 & & \\ -1.778 & -3.920 & 0.000 & -0.228 & -0.184 & -5.867 & \\ -1.778 & -3.920 & 0.000 & 0.228 & 0.184 & -1.652 & -5.867 \end{bmatrix} \quad (13.8)$$

Again, diagonal elements are bigger than off-diagonal elements because the $1/r$ operator acting on a basis function ν will ensure that the largest contribution to the overall integral will come from the nucleus k on which basis function ν resides. Unless μ also has significant amplitude around that nucleus, it will multiply the result by roughly zero and the whole integral will be small. Again, positive values can arise when two different functions are involved even though electrons in a single orbital must always be attracted to nuclei and thus diagonal elements must always be negative. Note that the p orbitals all have different nuclear attractions. That is because, although they all have the same attraction to the O nucleus, they have different amplitudes at the H nuclei. The p_x orbital has the smallest amplitude at the H nuclei (zero, since they are in its nodal plane), so it has the smallest nuclear attraction integral. The p_z orbital has somewhat smaller amplitude at the H nuclei than the p_y orbital because the bond angle is greater than 90° (it is 104.5° ; if it were 90° the O-H bonds would bisect the p_y and p_z orbitals and their amplitudes at the H nuclei would necessarily be the same). Thus, the nuclear attraction integral for the latter orbital is slightly smaller than for the former.

The sum of the kinetic and nuclear attraction integrals is usually called the one- electron or core part of the Fock matrix and abbreviated \mathbf{h} (i.e., $\mathbf{h} = \mathbf{K} + \mathbf{V}$). One then writes $\mathbf{F} = \mathbf{h} + \mathbf{G}$ where \mathbf{F} is the Fock matrix, \mathbf{h} is the one-electron matrix, and \mathbf{G} is the remaining part of the Fock matrix coming from the two-electron four-index integrals (cf eq. (13.4)). To compute those two-electron integrals, however, we need the density matrix, which itself comes from the occupied MO coefficients. So, we need an initial guess at

those coefficients. We can get such a guess many ways, but ultimately any guess is as good as any other. With these coefficients we can compute the density matrix using eq. (13.5):

$$\mathbf{P} = \begin{bmatrix} 2.108 \\ -0.456 & 2.010 \\ 0.000 & 0.000 & 2.000 \\ 0.000 & 0.000 & 0.000 & 0.737 \\ -0.104 & 0.618 & 0.000 & 0.000 & 1.215 \\ -0.022 & -0.059 & 0.000 & 0.539 & -0.482 & 0.606 \\ -0.022 & -0.059 & 0.000 & -0.539 & -0.482 & -0.183 & 0.606 \end{bmatrix} \quad (13.9)$$

With \mathbf{P} , we can compute the remaining contribution of \mathbf{G} to the Fock matrix. We will not list all 406 two-electron integrals here. Instead, we will simply write the total Fock matrix:

$$\mathbf{F} = \begin{bmatrix} -20.236 \\ -5.163 & -2.453 \\ 0.000 & 0.000 & -0.395 \\ 0.000 & 0.000 & 0.000 & -0.327 \\ 0.029 & 0.130 & 0.000 & 0.000 & -0.353 \\ -1.216 & -1.037 & 0.000 & -0.398 & 0.372 & -0.588 \\ -1.216 & -1.037 & 0.000 & 0.398 & 0.372 & -0.403 & -0.588 \end{bmatrix} \quad (13.10)$$

So, we're finally ready to solve the secular determinant, since we have \mathbf{F} and \mathbf{S} fully formed. When we do that, and then solve for the MO coefficients for each root E , we get new occupied MOs. Then, we iterate again, and

again, and again, until we are satisfied that further iterations will not change either our (i) energy, (ii) density matrix, or (iii) MO coefficients (it's up to the quantum chemist to decide what is considered satisfactory).

In our water calculation, if we monitor the energy at each step we find:

$E(RHF) = -74.893\ 002\ 803$	a.u. after 1 cycles	
$E(RHF) = -74.961\ 289\ 145$	a.u. after 2 cycles	
$E(RHF) = -74.961\ 707\ 247$	a.u. after 3 cycles	
$E(RHF) = -74.961\ 751\ 946$	a.u. after 4 cycles	(13.11)
$E(RHF) = -74.961\ 753\ 962$	a.u. after 5 cycles	
$E(RHF) = -74.961\ 754\ 063$	a.u. after 6 cycles	
$E(RHF) = -74.961\ 754\ 063$	a.u. after 7 cycles	

Which means that our original guess was really not too bad—off by a bit less than 0.1 a.u. or roughly 60 kcal mol⁻¹. Our guess energy is too high, as the variational principle guarantees that it must be. Our first iteration through the secular determinant picks up nearly 0.07 a.u., our next iteration an additional 0.000 42 or so, and by the end we are converged to within 1 nanohartree (0.000 0006 kcal mol⁻¹).

The final optimized MOs for water are:

	1	2	3	4	5	6	7
E	-20.24094	-1.27218	-0.62173	-0.45392	-0.39176	0.61293	0.75095
1	.99411	-.23251	.00000	-.10356	.00000	-.13340	.00000
2	.02672	.83085	.00000	.53920	.00000	.89746	.00000
3	.00000	.00000	.00000	.00000	1.0000	.00000	.00000
4	.00000	.00000	.60677	.00000	.00000	.00000	.99474
5	-.00442	-.13216	.00000	.77828	.00000	-.74288	.00000
6	-.00605	.15919	.44453	-.27494	.00000	-.80246	-.84542
7	-.00605	.15919	-.44453	-.27494	.00000	-.80246	.84542

(13.12)

where the first row reports the eigenvalues of each MO, in E_h (i.e., the energy of one electron in the MO). The sum of all of the occupied MO energies should be an underestimation of the total electronic energy because electron-electron repulsion will have been double counted. So, if we sum the occupied orbital energies (times two, since there are two electrons in each orbital), we get $2(-20.24094 - 1.27218 - 0.62173 - 0.45392 - 0.39176) = -45.961\ 060$. If we now subtract the electron-electron repulsion energy 38.265 406 we get -84.226 466. If we add the nuclear repulsion energy 9.264 701 to this we get a total energy -74.961 765. The difference between this and the converged result above (-74.961 754) can be attributed to rounding in the MO energies, which are truncated after 5 places. Notice that the five occupied MOs all have negative energies. So, their electrons are bound within the molecule. The unoccupied MOs (called “virtual” MOs) all have positive energies, meaning that the molecule will not spontaneously accept an electron from another source.

13.3 Shapes and Energies of Molecular Orbitals

If we analyze the optimized coefficients of the occupied MOs reported in eq. (13.12), we observe that the lowest energy orbital (by a lot!) is a nearly

pure oxygen 1s orbital since the coefficient of the oxygen 1s basis function is very nearly 1 and all other coefficients are rather close to 0. Note, however, that the coefficient is not really a percentage measure. That's because the basis functions are not necessarily orthogonal to one another. Let's consider the next molecular orbital up, number 2. It has a dominant contribution from the oxygen 2s basis function, but non-trivial contributions from many other basis functions as well. In order to understand which kind of orbital it is, it is useful to try to visualize some of its properties. For example, recall that the square of the orbital at a particular point in space represents a probability density. As such, we can map values of the square of each orbital on a grid in 3-dimensional space, and then pick a value of probability density, say $0.04 a_0^{-3}$, and plot that as a contour surface (remember that a probability density is a 4-dimensional quantity, so we need to take a slice at some constant density to be able to plot it in 3-D). That surface is called an “isodensity” surface. In addition to the square of the function, we can also color regions where the wave function is positive blue and regions where it's negative red. The five occupied and two unoccupied MOs mapped from their one-electron wave functions are plotted in figure 13.1.

Going back to the Lewis structure of water as taught in general chemistry courses, it says that there is one pair of electrons in one O–H σ bond, one pair in another identical such σ bond, and two equivalent pairs that constitute the lone pairs on oxygen. The two lone pairs and the O–H bonds should by pointing towards the apices of a tetrahedron because they are all considered to be sp^3 hybridized.

As you can see, the MOs look nothing like the Lewis picture. Instead, amongst other details, there is one lone pair that is pure p (not sp^3), another that is, if anything, sp^2 -like, but also enjoys contribution from hydrogen 1s components. There is one orbital that looks like both O–H σ bonds are present, but another that has an odd “bonding-all-over” character to it.

Is it really possible that for something as simple as water all the things you've ever been told about the Lewis structure are wrong? Water must have two equivalent lone pairs, right?

It turns out that the molecular orbital results can be tested with spectroscopic experiments, and suffice to say, they agree perfectly.

But the sp^3 -hybridized picture of water works well, for example, to explain its hydrogen-bonding behavior: In liquid water each water molecule makes two hydrogen bonds to other water molecules and accepts two more from different water molecules and the final structure has a net lattice-like form that is tetrahedral at each oxygen atom. How can the above MOs explain

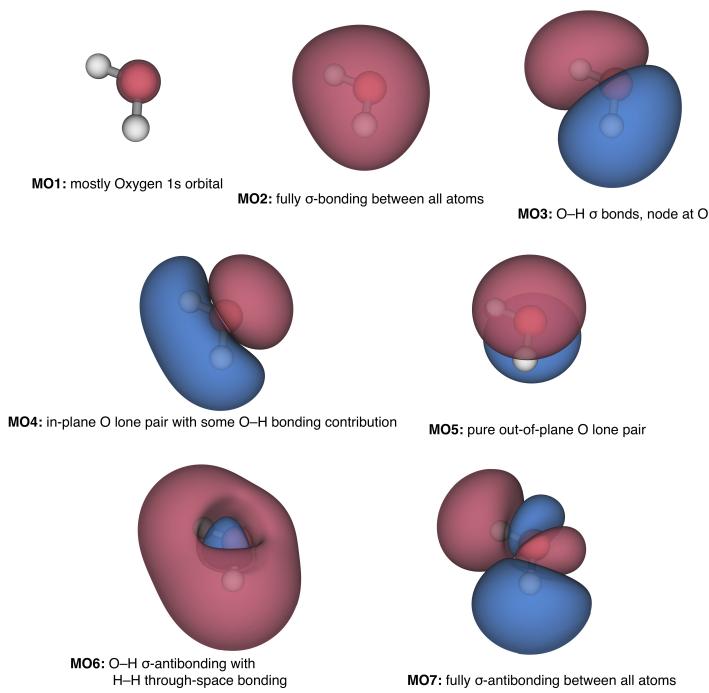


Figure 13.1: Isodensity maps of the Molecular Orbitals (MOs) of water.

that? The key point to remember is that another molecule does not see the individual orbitals of water, it just sees the final effect of all of those electrons and nuclei **together**. To explain the tetrahedral H-bond lattice we can plot some constant level of electron density (i.e. 0.02 a.u.) and map onto this isodensity surface the values of the electrostatic potential. We can find these values by bringing a positive test charge onto that surface and recording how much would it find itself attracted (because of a net negative electrostatic potential) or repelled (because of a net positive electrostatic potential). This is done in figure 13.2. Notice how the negative potential is entirely on the oxygen side and the positive potential entirely on the hydrogens side. Moreover, the negative potential splayes out to the tetrahedral points and the positive potential does too (those points for the purple region being roughly where the H atoms are).

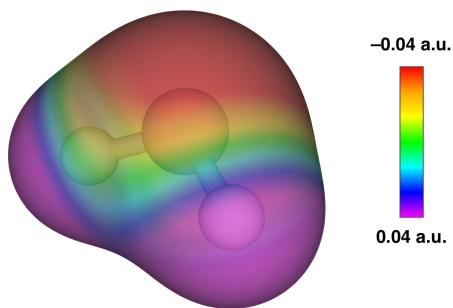


Figure 13.2: Molecular electrostatic potential of water plotted on a 0.02 a.u. isodensity surface.

13.4 Chapter Review

13.4.1 Study Questions

1. In the minimal-basis description of the water molecule, which atomic orbitals are included on oxygen?

- only 1s

- b. 1s and 2s
 - c. 2s and $2p_z$ only
 - d. 1s, 2s, and three 2p ($2p_x$, $2p_y$, $2p_z$)
 - e. only three 2p orbitals
-

2. How many atomic basis functions are used in the minimal representation of H₂O?

- a. 3
 - b. 5
 - c. 7
 - d. 9
 - e. 12
-

3. Which general method is used to build molecular orbitals for water from the chosen atomic basis in chapter 13?

- a. ure valence-bond resonance structures
 - b. LCAO-MO (linear combination of atomic orbitals to form molecular orbitals)
 - c. perturbation theory on hydrogenic orbitals only
 - d. classical hybridization rules only
 - e. direct diagonalization of the nuclear hamiltonian
-

4. In the minimal-basis HF treatment of water, the number of *one-electron* basis functions m determines the size of which matrix?

- a. the overlap and fock matrices, each of size $m \times m$
 - b. the nuclear–nuclear repulsion matrix of size $m^2 \times m^2$
 - c. a scalar energy value
 - d. the $3m \times 3m$ kinetic matrix
 - e. only the density matrix of size 2×2
-

5. In the minimal-basis water example, how many distinct one-electron integrals (kinetic or nuclear attraction) over contracted basis functions are there before considering primitive gaussians?

- a. 7
 - b. 49
 - c. 147
 - d. 196
 - e. 2,401
-

6. For the *two-electron* integrals over the 7 contracted basis functions on water, how many distinct contracted two-electron integrals are there (before primitives)?

- a. 7

- b. 49
 - c. 147
 - d. 196
 - e. 2,401
-

7. What main point does the water/minimal-basis example illustrate about *ab initio* calculations, even for small molecules?

- a. the number of two-electron integrals quickly becomes very large and dominates computational cost
 - b. only one-electron integrals matter for molecular properties
 - c. integrals scale linearly with system size
 - d. integral evaluation is trivial and negligible
 - e. minimal bases always give chemically exact results
-

8. In the minimal-basis H₂O example, which qualitative feature of the resulting MO diagram is reproduced correctly?

- a. exact experimental orbital energies
 - b. ordering of core, bonding, lone-pair, and antibonding orbitals
 - c. no distinction between bonding and nonbonding orbitals
 - d. completely delocalized electrons with no identifiable lone pairs
 - e. degenerate σ and π orbitals
-

9. What is the main physical interpretation of the two highest-occupied MOs in the water HF/minimal-basis picture?

- a. σ bonds along the H-O-H bisector
 - b. core orbitals on oxygen
 - c. primarily oxygen lone-pair orbitals oriented roughly toward the non-bonding directions
 - d. purely hydrogen 1s orbitals
 - e. totally delocalized over all three atoms with no directional character
-

10. Despite its simplicity, what important qualitative property of water does the minimal-basis HF treatment reproduce?

- a. linear geometry
 - b. no permanent dipole moment
 - c. equal O-H and H-H bond lengths
 - d. three-fold symmetry
 - e. bent geometry with approximate correct orientation of lone pairs and bonds
-

Answers: 1.d, 2.c, 3.b, 4.a, 5.d, 6.e, 7.a, 8.b, 9.c, 10.e

Chapter 14

Spectroscopy

The primary method of measuring the energy levels of a material is through the use of electromagnetic radiation. Experiments involving electromagnetic radiation—matter interaction are called spectroscopies. Since the energy levels of atoms and molecules are discontinuous, they absorb or emit light only at specific energies. These specific values correspond to the energy level difference between the initial and final states and they can be measured as signals in spectroscopic experiments. The intensity of the experimental signals depends on the population of the initial state involved in the transition.

Depending on the type of radiation, as well as the shape of the molecules and the inner details of the instrument that is used, some transition might be visible by the experiment (allowed), while others might not be (forbidden). The analysis of allowed and forbidden transition for each type of spectroscopy results into some mathematical formula that are called **selection rules**.

To summarize, spectroscopy is mainly the result of the following three effects:

- The energy levels of the atoms or molecules (determining the position of the signals).
- The population of the energy levels (determining the intensity of the signals).
- The selection rules that account for the symmetry and the interaction with the instrument.

Spectroscopy is the most important experimental verification of quantum mechanics, since we can use it to validate its theoretical results on the energy levels of atoms and molecules.

14.1 Rotational Spectroscopy

Rotational spectroscopy is concerned with the measurement of the energies of transitions between quantized rotational states of molecules in the gas phase. Rotational transitions of molecules are usually measured in the range $1 - 10 \text{ cm}^{-1}$ (microwave radiation) and rotational spectroscopy is therefore usually referred to as *microwave spectroscopy*.

Rotational spectroscopy is actively used by astrophysicists to explore the chemical composition of the interstellar medium using radio telescopes.

The rotational energies are derived theoretically by considering the molecules to be rigid rotors and applying the same treatment that we saw in chapter 4. Correction terms might be applied to account for deviation from the ideal rigid rotor case. As we saw in chapter 4, the quantized rotational energy levels of a rigid rotor depend on the angular moment of inertia, which in turn depends on the masses of the nuclei and the internuclear distance. Reversing the theoretical procedure of obtaining the energy levels from the distances, we can use the experimental energy levels to derive very precise values of molecular bond lengths (and in some complex case, also of angles). We will discuss below the simplest case of a diatomic molecule. For non-linear molecules, the moments of inertia are multiple, and only a few analytical method of solving the TISEq are available. For the most complicated cases, numerical methods can be used.

14.1.1 Rotation of diatomic molecules

Transitions between rotational states can be observed in molecules with a permanent electric dipole moment. The rigid rotor is a good starting point from which to construct a model of a rotating molecule. It is assumed that component atoms are point masses connected by rigid bonds. A linear molecule lies on a single axis and each atom moves on the surface of a sphere around the center of mass. The two degrees of rotational freedom correspond to the spherical coordinates, θ and φ , which describe the direction of the molecular axis. The quantum state is determined by two quantum numbers J and M . J defines the magnitude of the rotational angular momentum, and

M its component about an axis fixed in space, such as an external electric or magnetic field. In the absence of external fields, the energy depends only on J . Under the rigid rotor model, the rotational energy levels, $F(J)$, of the molecule can be expressed as:

$$F(J) = BJ(J+1) \quad J = 0, 1, 2, \dots \quad (14.1)$$

where B is the rotational constant of the molecule and is related to its moment of inertia. In a diatomic molecule the moment of inertia about an axis perpendicular to the molecular axis is unique, so:

$$B = \frac{\hbar^2}{8\pi^2 c I}, \quad (14.2)$$

with:

$$I = \frac{m_1 m_2}{m_1 + m_2} d^2, \quad (14.3)$$

where m_1 and m_2 are the masses of the atoms and d is the distance between them.

The selection rule for rotational spectroscopy dictate that during emission or absorption the rotational quantum number has to change by unity:

$$\Delta J = J' - J'' = \pm 1, \quad (14.4)$$

where J' denotes the lower level and J'' denotes the upper level involved in the transition. Thus, the locations of the lines in a rotational spectrum will be given by

$$\tilde{\nu}_{J' \leftrightarrow J''} = F(J') - F(J'') = 2B(J'' + 1) \quad J'' = 0, 1, 2, \dots \quad (14.5)$$

The diagram illustrates rotational transitions that obey the $\Delta J = 1$ selection rule is in figure 14.1.¹ The dashed lines show how these transitions map onto features that can be observed experimentally. Adjacent $J'' \leftarrow J'$ transitions are separated by $2B$ in the observed spectrum. Frequency or wavenumber units can also be used for the x axis of this plot.

¹This diagram is taken from Wikipedia by user Nnrw, and distributed under CC BY_SA 3.0 license.

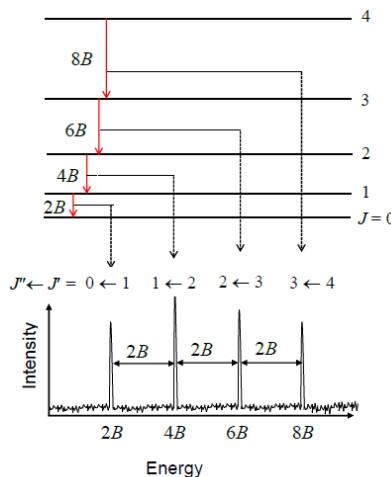


Figure 14.1: Rotational energy levels and line positions calculated in the rigid rotor approximation.

The probability of a transition taking place is the most important factor influencing the intensity of an observed rotational line. This probability is proportional to the population of the initial state involved in the transition. The population of a rotational state depends on two factors. The number of molecules in an excited state with quantum number J , relative to the number of molecules in the ground state, N_J/N_0 is given by the Boltzmann distribution:

$$\frac{N_J}{N_0} = e^{-\frac{E_J}{kT}} = \exp \left[-\frac{BhcJ(J+1)}{kT} \right], \quad (14.6)$$

where k is the Boltzmann constant and T is the absolute temperature. This factor decreases as J increases. The second factor is the degeneracy of the rotational state, which is equal to $2J+1$. This factor increases as J increases. Combining the two factors we obtain:

$$\text{population} \propto (2J+1) \exp \left[\frac{E_J}{kT} \right], \quad (14.7)$$

in agreement with the experimental shape of rotational spectra of diatomic molecules.

14.2 Vibrational Spectroscopy

Vibrational spectroscopy is concerned with the measurement of the energies of transitions between quantized vibrational states of molecules in the gas phase. These transitions usually occur in the middle infrared (IR) region of the electromagnetic wave at approximately $4,000 - 400 \text{ cm}^{-1}$ ($2.5 - 25 \mu\text{m}$). In the gas phase, vibrational transitions are almost always accompanied by changes in rotational energy. Transitions involving changes in both vibrational and rotational states are usually abbreviated as rovibrational transitions. Since changes in rotational energy levels are typically much smaller than changes in vibrational energy levels, changes in rotational state are said to give fine structure to the vibrational spectrum. For a given vibrational transition, the same theoretical treatment that we saw in the previous section for pure rotational spectroscopy gives the rotational quantum numbers, energy levels, and selection rules.

As we have done in the previous section, we will discuss below the simplest case of a diatomic molecule. For non-linear molecules the spectra becomes complicated to calculate, but their interpretation remains an important tool for the analysis of chemical structures.

14.2.1 Vibration of heteronuclear diatomic molecules

Diatomc molecules with the general formula AB have one normal mode of vibration involving stretching of the A – B bond. The vibrational term values, $G(v)$ can be calculated with the harmonic approximation that we discussed in chapter 4. The resulting equidistant energy levels depend on one vibrational quantum number v :

$$G(v) = \omega_e \left(v + \frac{1}{2} \right), \quad (14.8)$$

where ω_e is the harmonic frequency around equilibrium. When the molecule is in the gas phase, it can rotate about an axis, perpendicular to the molecular axis, passing through the center of mass of the molecule. As we discussed in the previous section, the rotational energy is also quantized, and depend on the rotational quantum number J . The values of the ro-vibrational states

are found (in wavenumbers) by combining the expressions for vibration and rotation:

$$G(v) + F_v(J) = \left[\omega_e \left(v + \frac{1}{2} \right) + B_v J(J+1) \right], \quad (14.9)$$

where $F_v(J)$ are the rotational levels at each vibrational state v .²

The selection rule for electric dipole allowed ro-vibrational transitions, in the case of a diamagnetic diatomic molecule is:

$$\Delta v = \pm 1 (\pm 2, \pm 3, \dots), \Delta J = \pm 1. \quad (14.10)$$

The transition with $\Delta v = \pm 1$ is known as the *fundamental transition*, while the others are called *overtones*. The selection rule has two consequences:

1. Both the vibrational and rotational quantum numbers must change. The transition $\Delta v = \pm 1, \Delta J = 0$ (Q-branch) is forbidden.
2. The energy change of rotation can be either subtracted from or added to the energy change of vibration, giving the P- and R- branches of the spectrum, respectively.

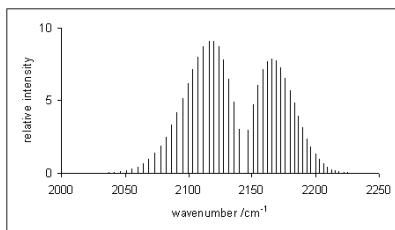


Figure 14.2: Simulated vibration-rotation line spectrum of carbon monoxide. The P-branch is to the left of the gap at 2140 cm^{-1} , the R-branch on the right.

A typical rovibrational spectrum is reported in figure 14.2 for the CO molecule.³ The intensity of the signals is—once again—proportional to the

²This is just a first approximation to rovibrational spectroscopy. Corrections for anharmonicity centrifugal distortion are necessary to closely match experimental spectra.

³This picture is taken from Wikipedia of anonymous user, and distributed under CC BY 3.0 license.

initial population of the levels. Notice how the signals in the spectrum are divided among two sides, the P-branch to the left, and the R-branch to the right. These signals correspond to the transitions reported in figure 14.3.⁴ Notice how the transitions corresponding to the Q-branch are forbidden by the selection rules, and therefore not observed in the experimental spectrum. The position of the missing Q-branch, however, can be easily obtained from the experimental spectrum as the missing signal between the P- and R- branches. Since the Q-branch transitions do not involve changes in the rotational energy level, their value is directly proportional to ω_e . This fact makes rovibrational spectroscopy an important experimental tool in the determination of bond distances of diatomic molecules.

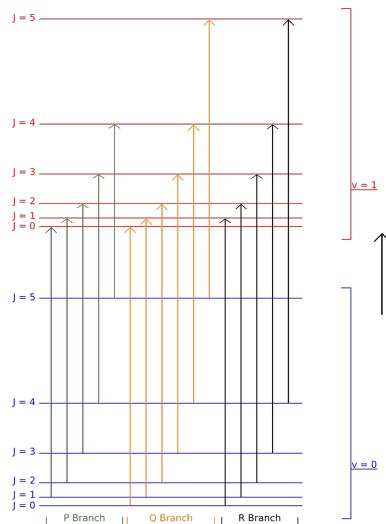


Figure 14.3: Schematic rovibrational energy level diagram for a linear molecule.

14.2.2 Vibration of homonuclear diatomic molecules

The quantum mechanics for homonuclear diatomic molecules is qualitatively the same as for heteronuclear diatomic molecules, but the selection rules governing transitions are different. Since the electric dipole moment of the

⁴This picture is taken from Wikipedia by user David-i98, and under public domain.

homonuclear diatomics is zero, the fundamental vibrational transition is electric-dipole-forbidden and the molecules are infrared inactive.

The spectra of these molecules can be observed by a type of IR spectroscopy that is subject to different selection rules. This technique is called **Raman spectroscopy**, and allows identification of the rovibrational spectra of homonuclear diatomic molecules because their molecular vibration is Raman-allowed.

14.3 Electronic Spectroscopy

Electronic spectroscopy is concerned with the measurement of the energies of transitions between quantized electronic states of molecules. Electronic transitions are always associated with simultaneous changes in vibrational levels. In the gas phase vibronic transitions are also accompanied by changes in rotational energy.

Electronic transitions are typically observed in the visible and ultraviolet regions, in the wavelength range approximately $200 - 700$ nm ($50,000 - 14,000\text{ cm}^{-1}$). When the electronic and vibrational energy changes are drastically different, vibronic coupling (mixing of electronic and vibrational wave functions) can be neglected and the energy of a vibronic level can be taken as the sum of the electronic and vibrational (and rotational) energies; that is, the Born–Oppenheimer approximation applies. The overall molecular energy depends not only on the electronic state but also on the vibrational and rotational quantum numbers, v and J . In this context, it is conventional to add a double prime (v'', J'') for levels of the electronic ground state and a single prime (v', J') for electronically excited states.

Each electronic transition may show vibrational coarse structure, and for molecules in the gas phase, rotational fine structure. This is true even when the molecule has a zero dipole moment and therefore has no vibration-rotation infrared spectrum or pure rotational microwave spectrum.

It is necessary to distinguish between absorption and emission spectra. With absorption the molecule starts in the ground electronic state, and usually also in the vibrational ground state $v'' = 0$ because at ordinary temperatures the energy necessary for vibrational excitation is large compared to the average thermal energy. The molecule is excited to another electronic state and to many possible vibrational states $v' = 0, 1, 2, 3, \dots$. With emission, the molecule can start in various populated vibrational states, and finishes in the electronic ground state in one of many populated vibrational levels.

The emission spectrum is more complicated than the absorption spectrum of the same molecule because there are more changes in vibrational energy level.

As we did for the previous two cases, we will concentrate below on the electronic absorption spectroscopy of diatomic molecules.

14.3.1 Electronic spectroscopy of diatomic molecules

The vibronic spectra of diatomic molecules in the gas phase also show rotational fine structure. Each line in a vibrational progression will show P- and R- branches. For some electronic transitions there will also be a Q-branch. The transition energies of the lines for a particular vibronic transition are given (in wavenumbers) by:

$$G(J', J'') = \bar{\nu}_{v'-v''} + B' J'(J' + 1) - B'' J''(J'' + 1). \quad (14.11)$$

The values of the rotational constants, B' and B'' may differ appreciably because the bond length in the electronic excited state may be quite different from the bond length in the ground state. The rotational constant is inversely proportional to the square of the bond length. Usually $B' < B''$, as is true when an electron is promoted from a bonding orbital to an anti-bonding orbital, causing bond lengthening.

The treatment of rotational fine structure of vibronic transitions is similar to the treatment of rotation-vibration transitions and differs principally in the fact that the ground and excited states correspond to two different electronic states as well as to two different vibrational levels. For the P-branch $J' = J'' - 1$, so that:

$$\begin{aligned} \bar{\nu}_P &= \bar{\nu}_{v'-v''} + B'(J'' - 1)J'' - B'' J''(J'' + 1) \\ &= \bar{\nu}_{v'-v''} - (B' + B'')J'' + (B' - B'')J''^2. \end{aligned} \quad (14.12)$$

Similarly, for the R-branch $J'' = J' - 1$, and:

$$\begin{aligned} \bar{\nu}_R &= \bar{\nu}_{v'-v''} + B' J'(J' + 1) - B'' J'(J' - 1) \\ &= \bar{\nu}_{v'-v''} + (B' + B'')J' + (B' - B'')J'^2. \end{aligned} \quad (14.13)$$

Thus, the wavenumbers of transitions in both P- and R- branches are given, to a first approximation, by the single formula:

$$\bar{\nu}_{P,R} = \bar{\nu}_{v',v''} + (B' + B'')m + (B' - B'')m^2, \quad m = \pm 1, \pm 2, \dots \quad (14.14)$$

Here positive m values refer to the R-branch (with $m = +J' = J'' + 1$) and negative values refer to the P-branch (with $m = -J''$).

The intensity of allowed vibronic transitions is governed by the Franck-Condon principle, which states that during an electronic transition, a change from one vibrational energy level to another will be more likely to happen if the two vibrational wave functions overlap more significantly. A diagrammatic representation of electronic spectroscopy and the Frack-Condon principle for a diatomic molecule is presented in figure 14.4.⁵

14.4 Chapter Review

14.4.1 Study Questions

1. In the rigid-rotor model for a diatomic molecule, the rotational energy levels are given by $F(J) = BJ(J+1)$. what does B represent?

- a. rotational constant related to the moment of inertia
 - b. bond length
 - c. rotational quantum number
 - d. vibrational frequency
 - e. electronic excitation energy
-

2. What are the allowed rotational quantum numbers J for a linear rigid rotor?

- a. $J = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$

⁵This picture is taken from Wikipedia by user Samoza, and distributed under CC BY-SA 3.0 license.

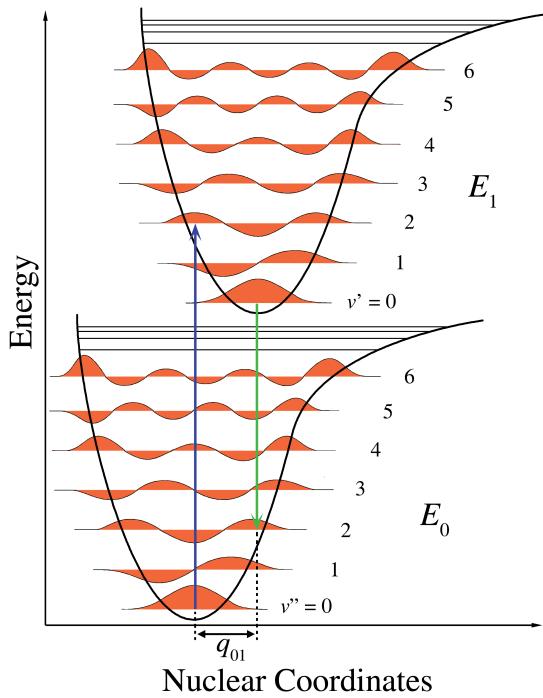


Figure 14.4: Energy level diagram illustrating the Franck–Condon principle.

- b. $J = 1, 2, 3, \dots$
 - c. $J = 0, 1, 2, 3, \dots$
 - d. $J = -\infty, \dots, \infty$
 - e. $J = 2, 4, 6, \dots$
-

3. In pure rotational spectroscopy of a heteronuclear diatomic in the rigid-rotor approximation, which of the following is the selection rule for rotational transitions?

- a. $\Delta J = 0$
 - b. $\Delta J = \pm 1$
 - c. $\Delta J = \pm 2$
 - d. $\Delta J = \pm 3$
 - e. no selection rule applies
-

4. For a harmonic oscillator model of a diatomic vibration, the vibrational energy levels are $G(v) = \omega_e \left(v + \frac{1}{2} \right)$. what are the allowed values of v ?

- a. $v = 0, 1, 2, \dots$
 - b. $v = 1, 2, 3, \dots$
 - c. $v = -\infty, \dots, \infty$
 - d. $v = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$
 - e. $v = 2, 4, 6, \dots$
-

5. What is the main vibrational selection rule for an infrared-active diatomic in the harmonic approximation?

- a. $\Delta v = 0$
 - b. $\Delta v = \pm 1$
 - c. $\Delta v = \pm 2$ only
 - d. $\Delta v = \pm 3$
 - e. no selection rule
-

6. In rovibrational spectroscopy of a diatomic molecule, transitions are often labeled as P- and R-branches. These correspond, respectively, to which of the following?

- a. $\Delta J = 0$ (P) and $\Delta J = \pm 2$ (R)
 - b. $\Delta J = -1$ (P) and $\Delta J = +1$ (R)
 - c. $\Delta J = +1$ (P) and $\Delta J = -1$ (R)
 - d. $\Delta J = 0$ for both P and R
 - e. $\Delta J = \pm 3$
-

7. Which of the following approximates the rovibrational transition wavenumbers in a simple harmonic-oscillator plus rigid-rotor model?

- a. a single line independent of j
- b. $\nu \approx \nu_0^2$
- c. $\nu \approx 2BJ(J+1)$ only

d. $\nu \approx \nu_0 \pm 2B(J + 1)$ for p and r branches

e. a random distribution

8. In electronic spectroscopy of diatomic molecules, which combination of changes typically occurs in an electronic transition?

a. electronic state changes, but vibrational and rotational states remain fixed

b. only rotational quantum number changes

c. primarily vibrational changes with no electronic change

d. electronic, vibrational, and rotational quantum numbers can all change, giving vibronic-rotational structure

e. only spin quantum numbers change

9. Why are electronic transitions typically found at much higher energies (shorter wavelengths) than vibrational or rotational transitions?

a. electronic masses are larger than nuclear masses

b. changes in electronic energy levels are usually much larger than nuclear vibrational or rotational spacings

c. vibrational transitions require relativistic corrections

d. rotational transitions do not conserve energy

e. electronic transitions have zero frequency

10. Which region is most commonly associated with *electronic* transitions in small molecules?

- a. microwave region
 - b. infrared region
 - c. visible and ultraviolet regions
 - d. radio-frequency region
 - e. far-infrared only
-

Answers: 1.a, 2.c, 3.b, 4.a, 5.b, 6.b, 7.d, 8.d, 9.b, 10.c