INTRODUCTION TO QUANTUM MECHANICS



Delmar Larsen
University of California, Davis

LibreTexts™

University of California, Davis

Chemistry 110A: Introduction to Quantum Mechanics

Delmar Larsen

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

This text was compiled on 03/11/2026

TABLE OF CONTENTS

Licensing

1: The Dawn of the Quantum Theory

- 1.1: Blackbody Radiation Cannot Be Explained Classically
- 1.2: Quantum Hypothesis used for Blackbody Radiation Law
- 1.3: Photoelectric Effect Explained with Quantum Hypothesis
- 1.4: The Hydrogen Atomic Spectrum
- 1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum
- 1.6: Matter Has Wavelike Properties
- 1.7: de Broglie Waves can be Experimentally Observed
- 1.8: The Bohr Theory of the Hydrogen Atom
- 1.9: The Heisenberg Uncertainty Principle
- 1.E: The Dawn of the Quantum Theory (Exercises)

2: The Classical Wave Equation

- 2.1: The One-Dimensional Wave Equation
- 2.2: The Method of Separation of Variables
- 2.3: Oscillatory Solutions to Differential Equations
- 2.4: The General Solution is a Superposition of Normal Modes
- 2.5: A Vibrating Membrane
- 2.E: The Classical Wave Equation (Exercises)

3: The Schrödinger Equation & a Particle in a Box

- 3.1: The Schrödinger Equation
- 3.2: Linear Operators in Quantum Mechanics
- 3.3: The Schrödinger Equation is an Eigenvalue Problem
- 3.4: The Quantum Mechanical Free Particle
- 3.5: Wavefunctions Have a Probabilistic Interpretation
- 3.6: The Energy of a Particle in a Box is Quantized
- 3.7: Wavefunctions Must Be Normalized
- 3.8: The Average Momentum of a Particle in a Box is Zero
- 3.9B: Particle in a Finite Box and Tunneling (optional)
- 3.9: The Uncertainty Principle Redux - Estimating Uncertainties from Wavefunctions
- 3.10: A Particle in a Two-Dimensional Box
- 3.11: A Particle in a Three-Dimensional Box
- 3.E: The Schrödinger Equation and a Particle in a Box (Exercises)

4: Postulates and Principles of Quantum Mechanics

- 4.1: The Wavefunction Specifies the State of a System
- 4.2: Quantum Operators Represent Classical Variables
- 4.3: Observable Quantities Must Be Eigenvalues of Quantum Mechanical Operators
- 4.4: The Time-Dependent Schrödinger Equation
- 4.5: The Eigenfunctions of Operators are Orthogonal
- 4.6: Heisenburg Uncertainy Principle III - Commuting Operators

- 4.E: Postulates and Principles of Quantum Mechanics (Exercises)

5: The Harmonic Oscillator and the Rigid Rotor

- 5.1: A Harmonic Oscillator Obeys Hooke's Law
- 5.2: The Equation for a Harmonic-Oscillator Model of a Diatomic Molecule Contains the Reduced Mass of the Molecule
- 5.3: The Harmonic Oscillator is an Approximation
- 5.4: The Harmonic Oscillator Energy Levels
- 5.5: The Harmonic Oscillator and Infrared Spectra
- 5.6: The Harmonic-Oscillator Wavefunctions Involve Hermite Polynomials
- 5.7: Hermite Polynomials are either Even or Odd Functions
- 5.8: The Energy Levels of a Rigid Rotor
- 5.9: The Rigid Rotator is a Model for a Rotating Diatomic Molecule
- 5.E: The Harmonic Oscillator and the Rigid Rotor (Exercises)

6: The Hydrogen Atom

- 6.1: The Schrödinger Equation for the Hydrogen Atom Can Be Solved Exactly
- 6.2: The Wavefunctions of a Rigid Rotator are Called Spherical Harmonics
- 6.3: The Three Components of Angular Momentum Cannot be Measured Simultaneously with Arbitrary Precision
- 6.4: Hydrogen Atomic Orbitals Depend upon Three Quantum Numbers
- 6.5: s Orbitals are Spherically Symmetric
- 6.6: Orbital Angular Momentum and the p-orbitals
- 6.7: The Helium Atom Cannot Be Solved Exactly
- 6.E: The Hydrogen Atom (Exercises)

7: Approximation Methods

- 7.0: Slater's Rules for Shielding
- 7.1: The Variational Method
- 7.2: Linear Variational Method and the Secular Determinant
- 7.3: Trial Functions Can Be Linear Combinations of Functions with Variational Parameters
- 7.4: Perturbation Theory Expresses the Solutions in Terms of Solved Problems
- 7.E: Approximation Methods (Exercises)

8: Multielectron Atoms

- 8.1: Atomic and Molecular Calculations are Expressed in Atomic Units
- 8.2: Perturbation Theory and the Variational Method for Helium
- 8.3: Hartree-Fock Equations are Solved by the Self-Consistent Field Method
- 8.4: An Electron Has An Intrinsic Spin Angular Momentum
- 8.5: Wavefunctions must be Antisymmetric to Interchange of any Two Electrons
- 8.6: Antisymmetric Wave Functions can be Represented by Slater Determinants
- 8.7: Hartree-Fock Calculations Give Good Agreement with Experimental Data
- 8.8: Term Symbols Gives Detailed Descriptions of an Electron Configuration
- 8.8B: Multi-electron Considerations - A Closer Look at Helium
- 8.9: The Allowed Values of J - the Total Angular Momentum Quantum Number
- 8.10: Hund's Rules Determine the Term Symbols of the Ground Electronic States
- 8.11: Using Atomic Term Symbols to Describe Atomic Spectra
- 8.E: Multielectron Atoms (Exercises)

9: The Chemical Bond: Diatomic Molecules

- 9.1: The Born-Oppenheimer Approximation Simplifies the Molecular Schrödinger Equation
- 9.2: The H_2^+ Prototypical Species
- 9.2B: Solving the H_2^+ System Exactly (Optional)
- 9.3: The Overlap Integral
- 9.3B: Evaluating the Overlap Integral (Optional)
- 9.4: Chemical Bond Stability
- 9.5: Bonding and Antibonding Orbitals
- 9.6: A Simple Molecular-Orbital Treatment of H_2 Places Both Electrons in a Bonding Orbital
- 9.7: Molecular Orbitals Can Be Ordered According to Their Energies
- 9.8: Molecular-Orbital Theory Does not Predict a Stable Diatomic Helium Molecule
- 9.9: Electrons Populate Molecular Orbitals According to the Pauli Exclusion Principle
- 9.10: Molecular Orbital Theory Predicts that Molecular Oxygen is Paramagnetic
- 9.E: The Chemical Bond: Diatomic Molecules (Exercises)

MathChapters

- A: Complex Numbers
- B: Probability and Statistics
- C: Vectors
- D: Spherical Coordinates
- E: Determinants
- F: Matrices
- G: Numerical Methods
- H: Partial Differentiation
- I: Series and Limits
- J: The Binomial Distribution and Stirling's Appromixation

Index

Glossary

Detailed Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

CHAPTER OVERVIEW

1: The Dawn of the Quantum Theory

"A profound change has taken place during the present century in the opinions physicists have held on the mathematical foundations of their subject. Previously they supposed that the principles of Newtonian mechanics would provide the basis for the description of the whole of physical phenomenon and that all the theoretical physicists had to do was suitably to develop and apply these principles. With the recognition that there is no logical reason why Newtonian and classical principles should be valid outside the domains in which they have been experimentally verified has come the realization that departures from these principles are indeed necessary. Such departures find their expression through the introduction of new mathematical formalisms, new schemes of axioms and rules of manipulation, into the methods of theoretical physics." **P. A. M. Dirac, "Quantum Mechanics" (1930).**

1.1: Blackbody Radiation Cannot Be Explained Classically
1.2: Quantum Hypothesis used for Blackbody Radiation Law
1.3: Photoelectric Effect Explained with Quantum Hypothesis
1.4: The Hydrogen Atomic Spectrum
1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum
1.6: Matter Has Wavelike Properties
1.7: de Broglie Waves can be Experimentally Observed
1.8: The Bohr Theory of the Hydrogen Atom
1.9: The Heisenberg Uncertainty Principle
1.E: The Dawn of the Quantum Theory (Exercises)

1: The Dawn of the Quantum Theory is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.1: Blackbody Radiation Cannot Be Explained Classically

Learning Objectives

One experimental phenomenon that could not be adequately explained by classical physics was blackbody radiation. Objectives for this section include

- Be familiar with black-body radiators
- Apply Stefan-Boltzmann's Law to estimate total light output from a radiator
- Apply Wien's Displacement Law to estimate the peak wavelength (or frequency) of the output from a black body radiator
- Understand the Rayleigh-Jeans Law and how it fails to properly model black-body radiation

All normal matter at temperatures above absolute zero emits electromagnetic radiation, which represents a conversion of a body's internal thermal energy into electromagnetic energy, and is therefore called *thermal radiation*. Conversely, all normal matter *absorbs* electromagnetic radiation to some degree. An object that absorbs ALL radiation falling on it, at all wavelengths, is called a blackbody. When a blackbody is at a uniform temperature, its emission has a characteristic frequency distribution that depends on the temperature. This emission is called *blackbody radiation*.

A room temperature blackbody appears black, as most of the energy it radiates is infra-red and cannot be perceived by the human eye. Because the human eye cannot perceive light waves at lower frequencies, a black body, viewed in the dark at the lowest just faintly visible temperature, subjectively appears grey, even though its objective physical spectrum peaks in the infrared range. When it becomes a little hotter, it appears dull red. As its temperature increases further it becomes yellow, white, and ultimately blue-white.



Figure 1.1.1 : Blackbody Radiation. When heated, all objects emit electromagnetic radiation whose wavelength (and color) depends on the temperature of the object. A relatively low-temperature object, such as a horseshoe forged by a blacksmith, appears red, whereas a higher-temperature object, such as the surface of the sun, appears yellow or white. Images used with permission from Wikipedia.

Blackbody radiation has a characteristic, continuous frequency spectrum that experimentally depends only on the body's temperature. In fact, we can be much more precise:

> *A body emits radiation at a given temperature and frequency exactly as well as it absorbs the same radiation.*

This statement was proved by Gustav Kirchhoff: the essential point is that if we *instead* suppose a particular body can absorb better than it emits, then in a room full of objects all at the same temperature, it will absorb radiation from the other bodies better than it radiates energy back to them. This means it will get hotter, and the rest of the room will grow colder, contradicting the second law of thermodynamics. *Thus, a body must emit radiation exactly as well as it absorbs the same radiation at a given temperature and frequency in order to not violate the second law of thermodynamics.*

Any body at any temperature above absolute zero **will** radiate to some extent, the intensity and frequency distribution of the radiation depending on the detailed structure of the body. To begin analyzing heat radiation, we need to be specific about the body doing the radiating: *the simplest possible case is an idealized body which is a perfect absorber, and therefore also (from the above argument) a perfect emitter.* So how do we construct a perfect absorber in the laboratory? In 1859 Kirchhoff had a good idea: a small hole in the side of a large box is an excellent absorber, since any radiation that goes through the hole bounces around inside, a

lot getting absorbed on each bounce, and has little chance of ever getting out again. So, we can do this *in reverse*: have an oven with a tiny hole in the side, and presumably the radiation coming out the hole is as good a representation of a perfect emitter as we're going to find (Figure 1.1.2).

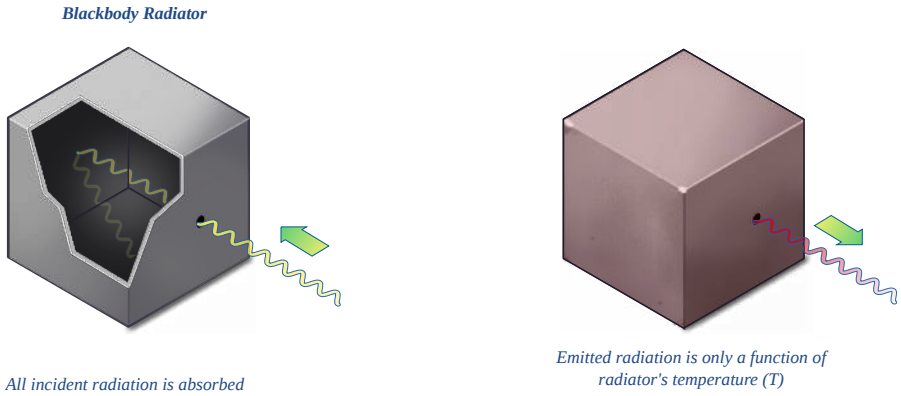


Figure 1.1.2 : Blackbody radiator is any object that is a perfect emitter and a perfect absorber of radiation. (CC BY-NC; Ümit Kaya via LibreTexts)

By the 1890's, experimental techniques had improved sufficiently that it was possible to make fairly precise measurements of the energy distribution of blackbody radiation. In 1895, at the University of Berlin, Wien and Lummer punched a small hole in the side of an otherwise completely closed oven, and began to measure the radiation coming out. The beam coming out of the hole was passed through a diffraction grating, which sent the different wavelengths/frequencies in different directions, all towards a screen. A detector was moved up and down along the screen to find how much radiant energy was being emitted in each frequency range. They found a radiation intensity/frequency curve close to the distributions in Figure 1.1.3 .

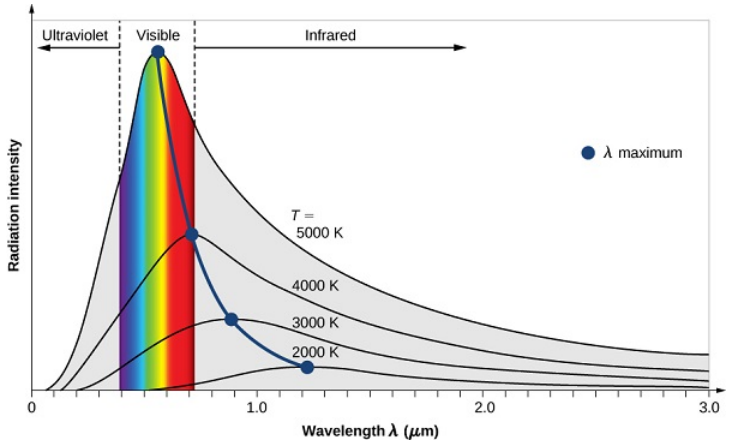
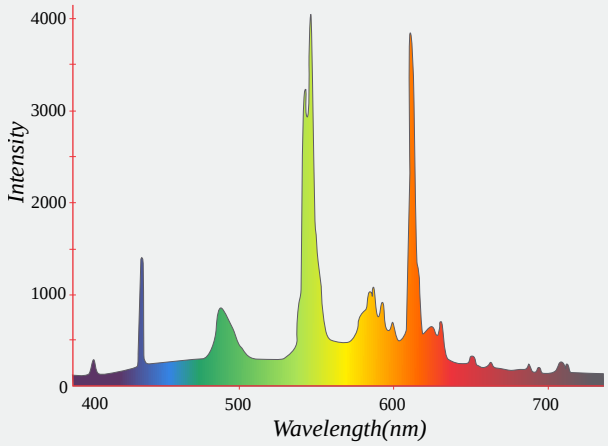


Figure 1.1.3 : Graphic representation of spectral distribution of blackbody radiation at different temperatures. The Stefan-Boltzmann's Law is observed as the increase in the emission amplitude with increasing temperature and the Wien's Displacement Law is observed as the shift to smaller wavelength with increasing temperature. (CC-BY 4.0; OpenStax)

By measuring the blackbody emission curves at different temperatures (Figure 1.1.3), they were also able to construct two important phenomenological Laws (i.e., formulated from experimental observations, not from basic principles of nature): **Stefan-Boltzmann's Law** and **Wien's Displacement Law**.

Not all radiators are blackbody radiators

The radiation of a blackbody radiator is produced by the thermal activity of the material, not the nature of the material, nor how it got thermally excited. Some examples of blackbodies include incandescent light bulbs, stars, and hot stove tops. The emission appears as a continuous spectrum (Figure 1.1.3) with multiple coexisting colors. However, not every radiator is a blackbody radiator. For example, the emission of a fluorescence bulb is not one. The following spectrum show the distribution of light from a fluoresce light tube and is a mixture of discrete bands at different wavelengths of light in contrast to the continuous spectra in Figure 1.1.3 for blackbody radiators.



Fluorescent lighting spectrum with emission peaks. Graph of Intensity (counts) vs. Wavelength (nm) in the visible spectrum. (CC BY-NC; Ümit Kaya via LibreTexts).

Fluorescent light bulbs contain a mixture of inert gases (usually argon and neon) together with a drop of mercury at low pressure. A different mix of visible colors blend to produce a light that appears to us white with different shadings.

The Stefan-Boltzmann Law

The first quantitative conjecture based on experimental observations was the **Stefan-Boltzmann Law** (1879) which states the total power (i.e., integrated over all emitting frequencies in Figure 1.1.3) radiated from one square meter of black surface goes as the *fourth power* of the absolute temperature (Figure 1.1.4):

$$P = \sigma T^4 \tag{1.1.1}$$

where

- P is the total amount of radiation emitted by an object per square meter ($Watts/m^2$)
- σ is a constant called the Stefan-Boltzman constant ($5.67 \times 10^{-8}\ Watts\ m^{-2}K^{-4}$)
- T is the absolute temperature of the object (in K)

The Stefan-Boltzmann Law is easily observed by comparing the integrated value (i.e., under the curves) of the experimental black-body radiation distribution in Figure 1.1.3 at different temperatures. In 1884, Boltzmann derived this T^4 behavior from theory by applying classical thermodynamic reasoning to a box filled with electromagnetic radiation, using Maxwell's equations to relate pressure to energy density. That is, the tiny amount of energy coming out of the hole (Figure 1.1.2) would of course have the same temperature dependence as the radiation intensity inside.

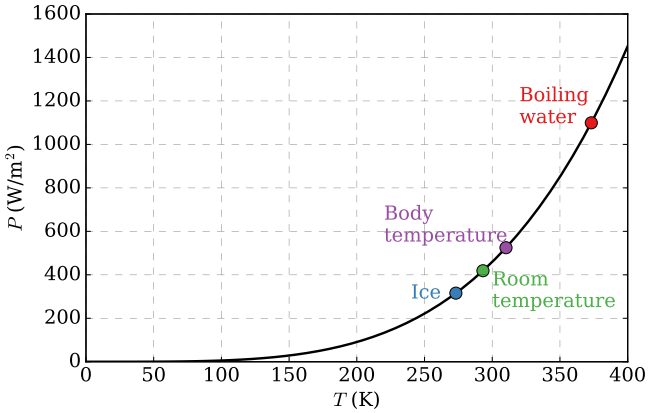


Figure 1.1.4 : Graph of a function of total emitted energy of a blackbody proportional to the fourth power of its thermodynamic temperature T according to the Stefan–Boltzmann law. (CC BY-SA 4.0; Nicoguaro)

Example 1.1.1

The sun's surface temperature is 5700 K.

a. How much power is radiated by the sun?
b. Given that the distance to earth is about 200 sun radii, what is the maximum power possible from a one square kilometer solar energy installation?

Solution

(a) First, we calculate the area of the sun followed by the flux (power). The sun has a radius of $6.96 \times 10^8 m$

The area of the sun is $A = 4\pi R^2$.

$$\begin{aligned} A &= 4(3.1416)(6.96 \times 10^8 m)^2 \\ &= 6.08 \times 10^{18} m^2 \end{aligned}$$

The power radiated from the sun (via Stefan-Boltzmann Law) is $P = \sigma T^4$ (Equation 1.1.1).

$$\begin{aligned} P &= (5.67 \times 10^{-8}\, Watts\, m^{-2} K^{-4})(5700K)^4 \\ &= 5.98 \times 10^7 Watts/m^2 \end{aligned}$$

This value is per square meter.

(b) To calculate the total power radiated by the sun is thus:

$$\begin{aligned} P_{total} &= PA = (5.98 \times 10^7 Watts/m^2)(6.08 \times 10^{18} m^2) \\ &= 3.6 \times 10^{26} Watts \end{aligned}$$

Wien's Displacement Law

The second phenomenological observation from experiment was **Wien's Displacement Law**. Wien's law identifies the dominant (peak) wavelength, or color, of light coming from a body at a given temperature. As the oven temperature varies, so does the frequency at which the emitted radiation is most intense (Figure 1.1.3). In fact, that frequency is directly proportional to the absolute temperature:

$$\nu_{max} \propto T \tag{1.1.2}$$

where the proportionality constant is $5.879 \times 10^{10} Hz/K$.

Wien himself deduced this law theoretically in 1893, following Boltzmann's thermodynamic reasoning. It had previously been observed, at least semi-quantitatively, by an American astronomer, Langley. This upward shift in ν_{max} with T is familiar to everyone—when an iron is heated in a fire (Figure 1.1.1), the first visible radiation (at around 900 K) is deep red, the lowest frequency visible light. Further increase in T causes the color to change to orange then yellow, and finally blue at very high temperatures (10,000 K or more) for which the peak in radiation intensity has moved beyond the visible into the ultraviolet.

Another representation of Wien's Law (Equation 1.1.2) in terms of the peak wavelength of light is

$$\lambda_{max} = \frac{b}{T} \tag{1.1.3}$$

where T is the absolute temperature in kelvin and b is a constant of proportionality called **Wien's displacement constant**, equal to $2.89 \times 10^{-3} m\ K$, or more conveniently to obtain wavelength in micrometers, $b \approx 2900\ \mu m \cdot K$. This is an inverse relationship between wavelength and temperature. So the higher the temperature, the shorter or smaller the wavelength of the thermal radiation. The lower the temperature, the longer or larger the wavelength of the thermal radiation. For visible radiation, hot objects emit bluer light than cool objects.

Example 1.1.2

If surface body temperature is 90 °F.

a. How much radiant energy in $W\, m^{-2}$ would your body emit?
b. What is the peak wavelength of emitted radiation?

c. What is the total radiant energy emitted by your body in Watts? Note: The average adult human male has a body surface area of about 1.9 m^2 and the average body surface area for a woman is about 1.6 m^2.

Solution

(a) 90 °F is 305 K. We use Stefan-Boltzmann Law (Equation 1.1.1). The total amount of radiation emitted will be $P = \sigma T^4$.

$$\begin{aligned} P &= (5.67 \times 10^{-8}\, Watts\, m^{-2} K^{-4})(305K)^4 \\ &= 491 W\, m^{-2} \end{aligned}$$

The peak wavelength of emitted radiation is found using Wien's Law:

$$\begin{aligned} \lambda_{max} &= \frac{2.898 \times 10^{-3} m \cdot K}{T} \\ &= \frac{2.898 \times 10^{-3} m \cdot K}{305K} \\ &= 9.5 \times 10^{-6} m = 9.5 \mu m \end{aligned}$$

The total radiant energy density in Watts is :

$$\begin{aligned} \text{Energy}_{\text{male}} &= (491 W\, m^{-2})(1.9 m^2) = 933 W \\ \text{Energy}_{\text{female}} &= (491 W\, m^{-2})(1.6 m^2) = 786 W \end{aligned}$$

Example 1.1.3 : The Temperature of the Sun

For example, if the Sun has a surface temperature of 5700 K, what is the wavelength of maximum intensity of solar radiation?

Solution

If we substitute 5700 K for T in Equation 1.1.3, we have

$$\begin{aligned} \lambda_{max} &= \frac{0.0029}{5700} \\ &= 5.1 \times 10^{-7}\, m \end{aligned}$$

Knowing that violet light has a wavelength of about 4.0×10^{-7} meters, yellow about 5.6×10^{-7} meters, and red about 6.6×10^{-7} meters, what can we say about the color of the Sun's peak radiation? The peak wavelength of the Sun's radiation is at a slightly shorter wavelength than the color yellow, so it is a slightly greenish yellow. To see this greenish tinge to the Sun, you would have to look at it from space. It turns out that the Earth's atmosphere scatters some of the shorter waves of sunlight, which shifts its peak wavelength to pure yellow.

Remember that thermal radiation always spans a wide range of wavelengths (Figure 1.1.2) and Equation 1.1.3 only specifies the single wavelength that is the **peak** of the spectrum. So although the Sun appears yellowish-white, when you disperse sunlight with a prism you see radiation with all the colors of the rainbow. Yellow just represents a characteristic wavelength of the emission.

Exercise 1.1.1

a. At what wavelength does the sun emit most of its radiation if it has a temperature of 5,778 K?
b. At what wavelength does the earth emit most of its radiation if it has a temperature of 288 K?

The Rayleigh-Jeans Law

Lord Rayleigh and J. H. Jeans developed an equation which explained blackbody radiation at low frequencies. The equation which seemed to express blackbody radiation was built upon all the known assumptions of physics at the time. *The big assumption which Rayleigh and Jean implied was that infinitesimal amounts of energy were continuously added to the system when the frequency was increased*. Classical physics assumed that energy emitted by atomic oscillations could have any continuous value. This was true for anything that had been studied up until that point, including things like acceleration, position, or energy. Their resulting **Rayleigh-Jeans Law** was

$$d\rho(\nu, T) = \rho_\nu(T)\, d\nu \tag{1.1.4}$$

$$= \frac{8\pi k_B T}{c^3}\nu^2 d\nu \tag{1.1.5}$$

Experimental data performed on the black box showed slightly different results than what was expected by the Rayleigh-Jeans law (Figure 1.1.5). The law had been studied and widely accepted by many physicists of the day, but the experimental results did not lie, something was different between what was theorized and what actually happens. The experimental results showed a bell type of curve, but according to the Rayleigh-Jeans law the frequency diverged as it neared the ultraviolet region (Equation 1.1.5). Ehrenfest later dubbed this the "ultraviolet catastrophe". It is important to emphasizing that Equation 1.1.5 is a **classical** result: the only inputs are classical dynamics and Maxwell's electromagnetic theory.

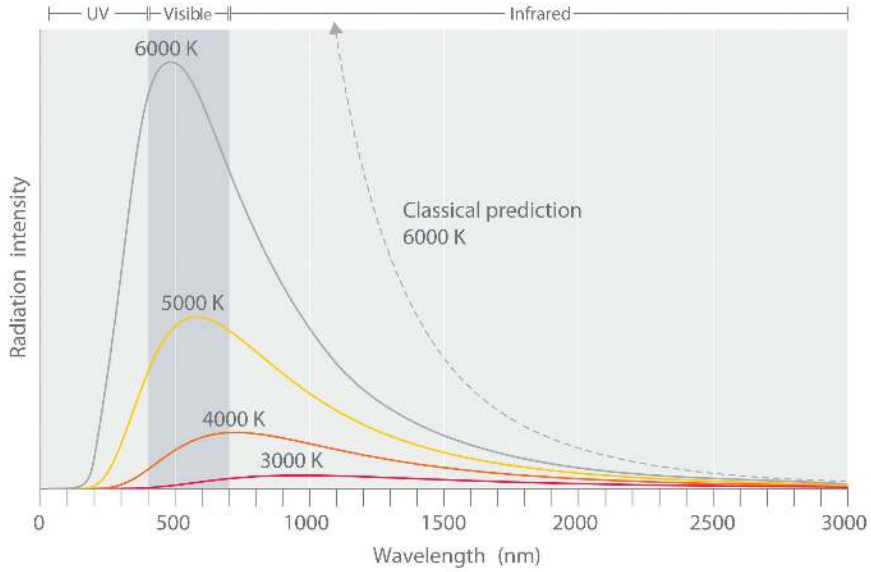


Figure 1.1.5 : Relationship between the temperature of an object and the spectrum of blackbody radiation it emits. At relatively low temperatures, most radiation is emitted at wavelengths longer than 700 nm, which is in the infrared portion of the spectrum. The dull red glow of the hot metalwork in Figure 1.1.5 is due to the small amount of radiation emitted at wavelengths less than 700 nm, which the eye can detect. As the temperature of the object increases, the maximum intensity shifts to shorter wavelengths, successively resulting in orange, yellow, and finally white light. At high temperatures, all wavelengths of visible light are emitted with approximately equal intensities. (CC BY-SA-NC; anonymous)

Differential vs. Integral Representation of the Distribution

Radiation is understood as a continuous distribution of amplitude vs. wavelength or, equivalently, amplitude vs. frequency (Figure 1.1.5). According to Rayleigh-Jeans law, the intensity at a specific frequency ν and temperature is

$$\rho(\nu, T) = \frac{8\pi k_b T \nu^2}{c^3}.$$

However, in practice, we are more interested in frequency intervals. An exact frequency is the limit of a sequence of smaller and smaller intervals. If we make the assumption that, for a sufficiently small interval, $\rho(\nu, T)$ does not vary, we get your definition for the differential $d\rho(\nu, T)$ in Equation 1.1.5:

The assumption is fair due to the continuity of $\rho(\nu, T)$. This is the approximation of an integral on a very small interval $d\nu$ by the height of a point inside this interval ($\frac{8\pi k_b T \nu^2}{c^3}$) times its length ($d\nu$). So, if we sum an infinite amount of small intervals like the one above we get an integral. The total radiation between ν_1 and ν_2 will be:

$$\int_{\nu_1}^{\nu_2} \mathrm{d}\rho(\nu,T) = \int_{\nu_1}^{\nu_2} \rho(\nu,T)\,\mathrm{d}\nu$$
$$= \int_{\nu_1}^{\nu_2} \frac{8\pi k_b T \nu^2}{c^3}\mathrm{d}\nu$$
$$= 8\pi k_b T \frac{v_2^3 - v_1^3}{3c^3}.$$

Observe that $\rho(\nu,T)$ is quadratic in ν.

Example 1.1.4 : The Ultraviolet Catastrophe

What is the total spectral radiance of a radiator that follows the Rayleigh-Jeans law for its emission spectrum?

Solution

The total spectral radiance $\rho_{tot}(T)$ is the combined emission over all possible wavelengths (or equivalently, frequencies), which is an integral over the relevant distribution (Equation 1.1.5 for the Rayleigh-Jeans Law).

$$\rho_{tot}(T) = \int_0^\infty d\rho\,(\nu,T)$$
$$= \int_0^\infty \frac{8\pi k_B T}{c^3}\nu^2 d\nu$$

but the integral

$$\int_0^\infty x^2 \mathrm{d}x$$

does not converge. Worse, it is infinite,

$$\lim_{k\to\infty} \int_0^k x^2 \mathrm{d}x = \infty$$

Hence, the classically derived Rayleigh-Jeans law predicts that the radiance of a blackbody is **infinite**. Since radiance is power per angle and unit area, this also implies that the total power and hence the energy a blackbody emitter gives off is infinite, which is patently absurd. This is called the ultraviolet catastrophe because the absurd prediction is caused by the classical law not predicting the behavior at high frequencies/small wavelengths correctly (Figure 1.1.5).

Contributors and Attributions

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)
- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")
- Paul Flowers (University of North Carolina - Pembroke), Klaus Theopold (University of Delaware) and Richard Langley (Stephen F. Austin State University) with contributing authors. Textbook content produced by OpenStax College is licensed under a Creative Commons Attribution License 4.0 license. Download for free at http://cnx.org/contents/85abf193-2bd...a7ac8df6@9.110).
- ACuriousMind (StackExchange)

1.1: Blackbody Radiation Cannot Be Explained Classically is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.1: Blackbody Radiation Cannot Be Explained Classically** is licensed CC BY 4.0.

1.2: Quantum Hypothesis used for Blackbody Radiation Law

Learning Objectives

- To understand how energy is quantized in blackbody radiation

By the late 19th century, many physicists thought their discipline was well on the way to explaining most natural phenomena. They could calculate the motions of material objects using Newton's laws of classical mechanics, and they could describe the properties of radiant energy using mathematical relationships known as Maxwell's equations, developed in 1873 by James Clerk Maxwell, a Scottish physicist. The universe appeared to be a simple and orderly place, containing matter, which consisted of particles that had mass and whose location and motion could be accurately described, and electromagnetic radiation, which was viewed as having no mass and whose exact position in space could not be fixed. Thus matter and energy were considered distinct and unrelated phenomena. Soon, however, scientists began to look more closely at a few inconvenient phenomena that could not be explained by the theories available at the time.

One experimental phenomenon that could not be adequately explained by classical physics was blackbody radiation (Figure 1.2.1). Attempts to explain or calculate this spectral distribution from classical theory were complete failures. A theory developed by Rayleigh and Jeans predicted that the intensity should go to infinity at short wavelengths. Since the intensity actually drops to zero at short wavelengths, the Rayleigh-Jeans result was called the **ultraviolet catastrophe** (Figure 1.2.1 dashed line). There was no agreement between theory and experiment in the ultraviolet region of the blackbody spectrum.

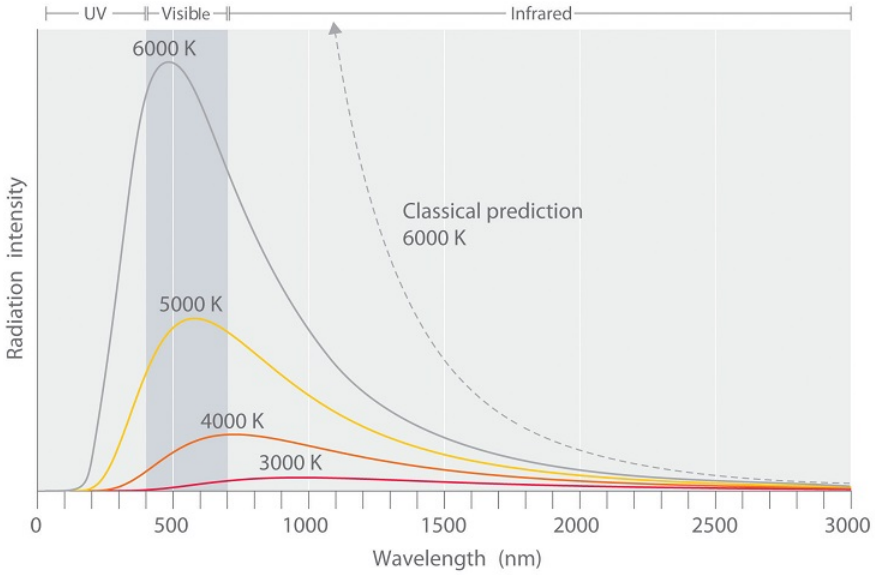


Figure 1.2.1 : Relationship between the temperature of an object and the spectrum of blackbody radiation it emits. At relatively low temperatures, most radiation is emitted at wavelengths longer than 700 nm, which is in the infrared portion of the spectrum. As the temperature of the object increases, the maximum intensity shifts to shorter wavelengths, successively resulting in orange, yellow, and finally white light. At high temperatures, all wavelengths of visible light are emitted with approximately equal intensities. The white light spectrum shown for an object at 6000 K closely approximates the spectrum of light emitted by the sun. Note the sharp decrease in the intensity of radiation emitted at wavelengths below 400 nm, which constituted the ultraviolet catastrophe. The classical prediction fails to fit the experimental curves entirely and does not have a maximum intensity. (CC BY-SA-NC; anonymous).

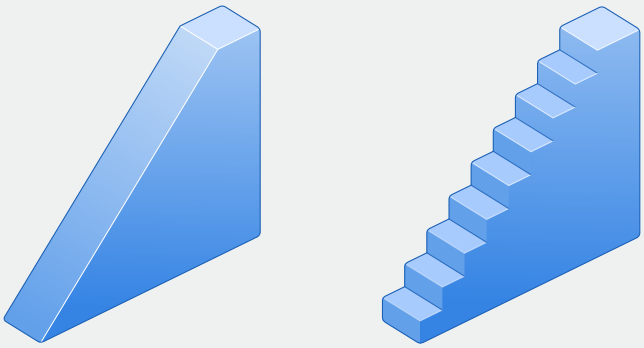
Quantizing Electrons in the Radiator

In 1900, the German physicist Max Planck (1858–1947) explained the ultraviolet catastrophe by proposing that the energy of electromagnetic waves is *quantized* rather than continuous. This means that for each temperature, there is a maximum intensity of radiation that is emitted in a blackbody object, corresponding to the peaks in Figure 1.2.1 , so the intensity does not follow a smooth curve as the temperature increases, as predicted by classical physics. Thus energy could be gained or lost only in integral

multiples of some smallest unit of energy, a quantum (the smallest possible unit of energy). Energy can be gained or lost only in integral multiples of a quantum.

Quantization

Although quantization may seem to be an unfamiliar concept, we encounter it frequently in quantum mechanics (hence the name). For example, US money is integral multiples of pennies. Similarly, musical instruments like a piano or a trumpet can produce only certain musical notes, such as C or F sharp. Because these instruments cannot produce a continuous range of frequencies, their frequencies are quantized. It is also similar to going up and down a hill using discrete stair steps rather than being able to move up and down a continuous slope. Your potential energy takes on discrete values as you move from step to step. Even electrical charge is quantized: an ion may have a charge of −1 or −2, but *not* −1.33 electron charges.



A continuous vs. a quantized (gravitationaly) potential energy system. In the continuous case (left) a system can have any potential energy, but in the quantized case (right), a system can only have certain values (other values are not allowed). (CC BY-NC; Ümit Kaya via LibreTexts)

Planck's quantization of energy is described by the his famous equation:

$$E = h\nu \tag{1.2.1}$$

where the proportionality constant h is called **Planck's constant**, one of the most accurately known fundamental constants in science

$$h = 6.626070040(81) \times 10^{-34}\, J \cdot s$$

However, for our purposes, its value to four significant figures is sufficient:

$$h = 6.626 \times 10^{-34}\, J \cdot s$$

As the frequency of electromagnetic radiation increases, the magnitude of the associated quantum of radiant energy increases. By assuming that energy can be emitted by an object only in integral multiples of $h\nu$, Planck devised an equation that fit the experimental data shown in Figure 1.2.1 . We can understand Planck's explanation of the ultraviolet catastrophe qualitatively as follows: At low temperatures, radiation with only relatively low frequencies is emitted, corresponding to low-energy quanta. As the temperature of an object increases, there is an increased probability of emitting radiation with higher frequencies, corresponding to higher-energy quanta. At any temperature, however, it is simply more probable for an object to lose energy by emitting a large number of lower-energy quanta than a single very high-energy quantum that corresponds to ultraviolet radiation. The result is a maximum in the plot of intensity of emitted radiation versus wavelength, as shown in Figure 1.2.1 , and a shift in the position of the maximum to lower wavelength (higher frequency) with increasing temperature.

At the time he proposed his radical hypothesis, Planck could not explain *why* energies should be quantized. Initially, his hypothesis explained only one set of experimental data—blackbody radiation. If quantization were observed for a large number of different phenomena, then quantization would become a law. In time, a theory might be developed to explain that law. As things turned out, Planck's hypothesis was the seed from which modern physics grew.

Max Planck explain the spectral distribution of blackbody radiation as result from oscillations of electrons. Similarly, oscillations of electrons in an antenna produce radio waves. Max Planck concentrated on modeling the oscillating charges that must exist in the oven walls, radiating heat inwards and—in thermodynamic equilibrium—themselves being driven by the radiation field. He found he could account for the observed curve if he required these oscillators not to radiate energy continuously, as the classical theory would demand, but they could **only** lose or gain energy in chunks, called **quanta**, of size $h\nu$, for an oscillator of frequency ν (Equation 1.2.1).

With that assumption, Planck calculated the following formula for the radiation energy density inside the oven:

$$d\rho(\nu,T) = \rho_\nu(T)d\nu \tag{1.2.2}$$

$$= \frac{8\pi h\nu^3}{c^3} \cdot \frac{1}{\exp\left(\frac{h\nu}{k_BT}\right)-1}d\nu \tag{1.2.3}$$

with

- $\pi = 3.14159$
- $h = 6.626 \times 10^{-34} J \cdot s$
- $c = 3.00 \times 10^8 m/s$
- $\nu = 1/s$
- $k_B = 1.38 \times 10^{-23} J/K$
- T is absolute temperature (in Kelvin)

Planck's radiation energy density (Equation 1.2.3) can also be expressed in terms of wavelength λ.

$$\rho(\lambda,T) = \frac{8\pi hc}{\lambda^5}\left(\frac{1}{e^{hc/\lambda k_BT}-1}\right) \tag{1.2.4}$$

With a wavelength of maximum energy density at:

$$\lambda_{max} = \frac{hc}{4.965kT}$$

This can be used to calculate the radiance emitted from surface per unit frequency

$$dB_\nu(\nu,T) = \rho_\nu(T)d\nu$$

$$= \frac{2h\nu^3}{c^2} \cdot \frac{1}{\exp\left(\frac{h\nu}{k_BT}\right)-1}d\nu \tag{1.2.5}$$

or

$$B_\lambda(\lambda,T) = \frac{2hc^2}{\lambda^5}\frac{1}{e^{hc/(\lambda k_BT)}-1} \tag{1.2.6}$$

Planck's equation (Equation 1.2.4) gave an excellent agreement with the experimental observations for all temperatures (Figure 1.2.2).

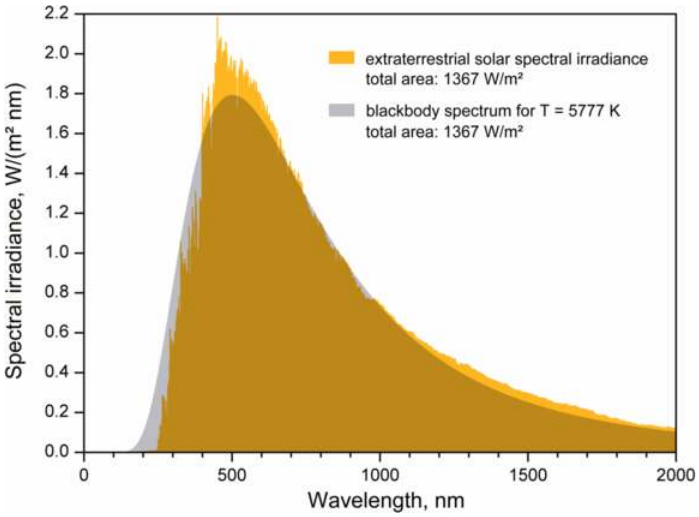
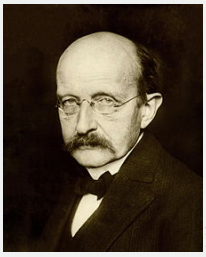


Figure 1.2.2 : The Sun is an excellent approximation of a blackbody. Its effective temperature is ~5777 K. (CC-SA-BY 3.0; Sch).

Max Planck (1858–1947)

Planck made many substantial contributions to theoretical physics, but his fame as a physicist rests primarily on his role as the originator of quantum theory. In addition to being a physicist, Planck was a gifted pianist, who at one time considered music as a career. During the 1930s, Planck felt it was his duty to remain in Germany, despite his open opposition to the policies of the Nazi government.



MAX-PLANCK-GESELLSCHAFT

(left) The German physicist Max Planck had a major influence on the early development of quantum mechanics, being the first to recognize that energy is sometimes quantized. Planck also made important contributions to special relativity and classical physics. (Public Domain; Library of Congress via Wikimedia) (left) The society's logo features Minerva, the Roman goddess of wisdom. (Fair use)

One of his sons was executed in 1944 for his part in an unsuccessful attempt to assassinate Hitler and bombing during the last weeks of World War II destroyed Planck's home. After WWII, the major German scientific research organization was renamed the Max Planck Society.

Exercise 1.2.1

Use Equation 1.2.4 to show that the units of $\rho(\lambda, T)\, d\lambda$ are J/m^3 as expected for an energy density.

The near perfect agreement of this formula with precise experiments (e.g., Figure 1.2.3), and the consequent necessity of energy quantization, was the most important advance in physics in the century. His blackbody curve was completely accepted as the correct one: more and more accurate experiments confirmed it time and again, yet the radical nature of the quantum assumption did not sink in. Planck was not too upset—he didn't believe it either, he saw it as a technical fix that (he hoped) would eventually prove unnecessary.

> *Part of the problem was that Planck's route to the formula was long, difficult and implausible—he even made contradictory assumptions at different stages, as Einstein pointed out later. However, the result was correct anyway!*

The mathematics implied that the energy given off by a blackbody was not continuous, but given off at certain specific wavelengths, in regular increments. If Planck assumed that the energy of blackbody radiation was in the form

$$E = nh\nu$$

where n is an integer, then he could explain what the mathematics represented. This was indeed difficult for Planck to accept, because at the time, there was no reason to presume that the energy should only be radiated at specific frequencies. Nothing in Maxwell's laws suggested such a thing. It was as if the vibrations of a mass on the end of a spring could only occur at specific energies. Imagine the mass slowly coming to rest due to friction, but not in a continuous manner. Instead, the mass jumps from one fixed quantity of energy to another without passing through the intermediate energies.

> *To use a different analogy, it is as if what we had always imagined as smooth inclined planes were, in fact, a series of closely spaced steps that only presented the illusion of continuity.*

Summary

The agreement between Planck's theory and the experimental observation provided strong evidence that the energy of electron motion in matter is quantized. In the next two sections, we will see that the energy carried by light also is quantized in units of $h\bar{\nu}$. These packets of energy are called "photons."

Contributors and Attributions

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)
- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")

1.2: Quantum Hypothesis used for Blackbody Radiation Law is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.2: Quantum Hypothesis Used for Blackbody Radiation Law** is licensed CC BY-NC-SA 4.0.

1.3: Photoelectric Effect Explained with Quantum Hypothesis

Learning Objectives

- To be familiar with the photoelectron effect for bulk materials
- Understand how the photoelectron kinetic energy and intensity vary as a function of incident light wavelength
- Understand how the photoelectron kinetic energy and intensity vary as a function of incident light intensity
- Describe what a workfunction is and relate it to ionization energy
- Describe the photoelectric effect with Einstein's quantized photon model of light

Nature, it seemed, was quantized (non-continuous, or discrete). If this was so, how could Maxwell's equations correctly predict the result of the blackbody radiator? Planck spent a good deal of time attempting to reconcile the behavior of electromagnetic waves with the discrete nature of the blackbody radiation, to no avail. It was not until 1905, with yet another paper published by Albert Einstein, that the wave nature of light was expanded to include the particle interpretation of light which adequately explained Planck's equation.

The photoelectric effect was first documented in 1887 by the German physicist Heinrich Hertz and is therefore sometimes referred to as the Hertz effect. While working with a spark-gap transmitter (a primitive radio-broadcasting device), Hertz discovered that upon absorption of certain frequencies of light, substances would give off a visible spark. In 1899, this spark was identified as light-excited electrons (called **photoelectrons**) leaving the metal's surface by J.J. Thomson (Figure 1.3.1).

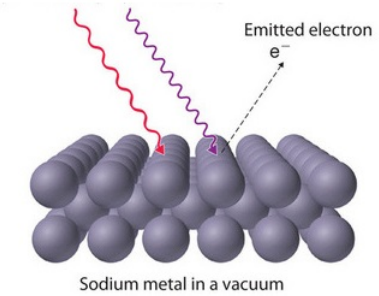


Figure 1.3.1 : The photoelectric effect involves irradiating a metal surface with photons of sufficiently high energy to cause the electrons to be ejected from the metal. (CC BY-SA-NC; anonymous)

The classical picture underlying the photoelectron effect was that the atoms in the metal contained electrons, that were shaken and caused to vibrate by the oscillating electric field of the incident radiation. Eventually some of them would be shaken loose, and would be ejected from the cathode. It is worthwhile considering carefully how the *number* and *speed* of electrons emitted would be expected to vary with the *intensity* and *color* of the incident radiation along with the time needed to observe the photoelectrons.

- Increasing the intensity of radiation would shake the electrons more violently, so one would expect more to be emitted, and they would shoot out at greater speed, on average.
- Increasing the frequency of the radiation would shake the electrons faster, so it might cause the electrons to come out faster. For very dim light, it would take some time for an electron to work up to a sufficient amplitude of vibration to shake loose.

Lenard's Experimental Results (Intensity Dependence)

In 1902, Hertz's student, Philipp Lenard, studied how the energy of the emitted photoelectrons varied with the intensity of the light. He used a carbon arc light and could increase the intensity a thousand-fold. The ejected electrons hit another metal plate, the collector, which was connected to the cathode by a wire with a sensitive ammeter, to measure the current produced by the illumination (Figure 1.3.2). To measure the energy of the ejected electrons, Lenard charged the collector plate negatively, to repel the electrons coming towards it. Thus, only electrons ejected with enough kinetic energy to get up this potential hill would contribute to the current.

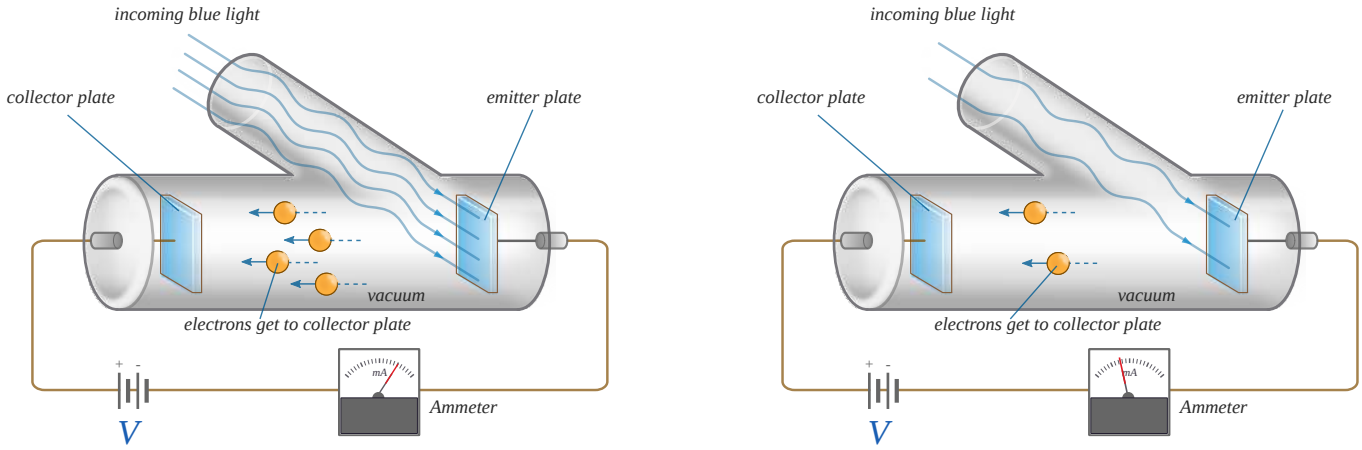


Figure **1.3.2**: Lenard's photoelectric experiment. (left) High light intensity increase photocurrent (number of collected photoelectrons). (right) Low light intensity has reduced photocurrent. However, the kinetic energy of the ejected electrons is independent of incident light intensity. (CC BY-NC; Ümit Kaya via LibreTexts)

Lenard discovered that there was a well defined minimum voltage that stopped any electrons getting through (V_{stop}). To Lenard's surprise, he found that V_{stop} did not depend at all on the intensity of the light! Doubling the light intensity doubled the *number* of electrons emitted, but did not affect the *kinetic energies* of the emitted electrons. The more powerful oscillating field ejected more electrons, but the maximum individual energy of the ejected electrons was the same as for the weaker field (Figure 1.3.2).

Millikan's Experimental Results (Wavelength Dependence)

The American experimental physicist Robert Millikan followed up on Lenard's experiments and using a powerful arc lamp, he was able to generate sufficient light intensity to separate out the colors and check the photoelectric effect using light of different colors. He found that the maximum energy of the ejected electrons *did* depend on the color - the shorter wavelength, higher frequency light eject photoelectrons with greater kinetic energy (Figures 1.3.3).

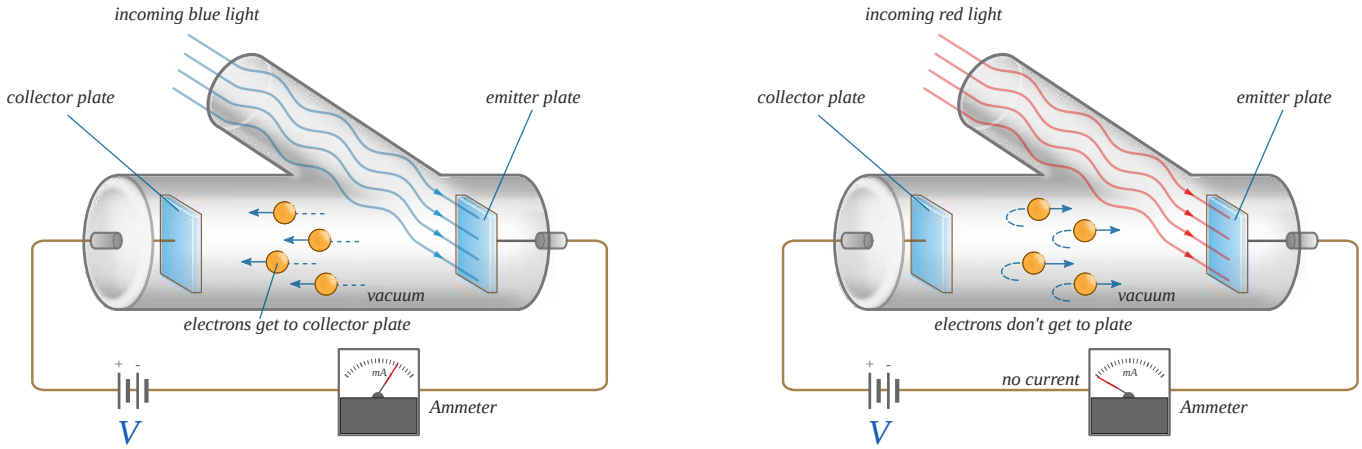


Figure 1.3.3 : Millikan's photoelectric experiment. (left) Incident high-energy blue light. The battery represents the potential Lenard used to charge the collector plate negatively, which would actually be a variable voltage source. Since the electrons ejected by the blue light are getting to the collector plate, the potential supplied by the battery is less than V_{stop}, for blue light. (right) Indicent low-energy red light. Since the electrons ejected by the red light are not getting to the collector plate, the potential supplied by the battery exceeds V_{stop} for red light. (CC BY-NC; Ümit Kaya via LibreTexts)

As shown in Figure 1.3.4 , just the opposite behavior from classical is observed from Lenard's and Millikan's experiments. The intensity affects the number of electrons, and the frequency affects the kinetic energy of the emitted electrons. From these sketches, we see that

- the kinetic energy of the electrons is linearly proportional to the frequency of the incident radiation above a threshold value of ν_0 (no current is observed below ν_0), and the kinetic energy is independent of the intensity of the radiation, and
- the number of electrons (i.e. the electric current) is proportional to the intensity and independent of the frequency of the incident radiation above the threshold value of ν_0 (i.e., no current is observed below ν_0).

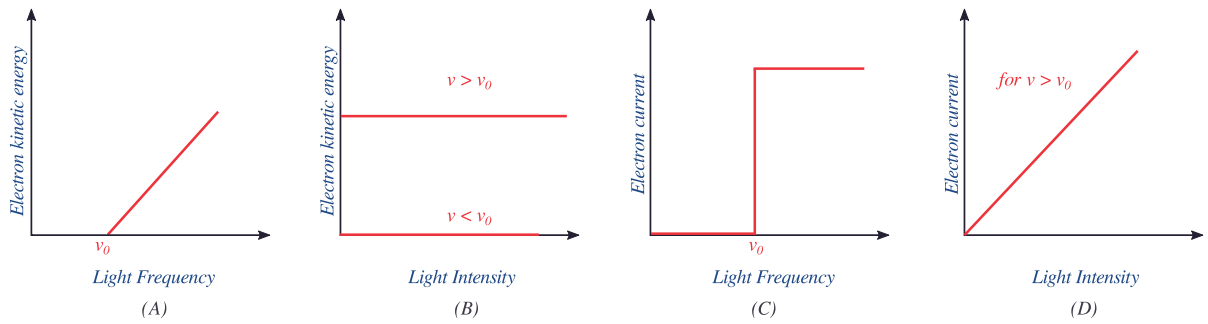


Figure 1.3.4 : Schematic drawings showing the characteristics of the photoelectric effect from Lenard's and Millikan's experiments. (A) The kinetic energy of any single emitted electron increases linearly with frequency above some threshold value (B) The electron kinetic energy is independent of the light intensity above the threshold frequency and zero below. (C) The number of electrons emitted per second (i.e. the electric current) is independent of light frequency above the threshhold frequency and zero below. (D) The number of electrons increases linearly with the light intensity. (CC BY-NC; Ümit Kaya via LibreTexts)

Classical Theory does not Describe Experiment

Classical theory predicts that energy carried by light is proportional to its amplitude independent of its frequency, and this fails to correctly explain the observed wavelength dependence in Lenard's and Millikan's observations.

As with most of the experimental results we discuss in this text, the behavior described above is a simplification of the true experimental results observed in the laboratory. A more complex description involves a greater introduction of more complex physics and instrumentation, which will be ignored for now.

Einstein's Quantum Picture

In 1905 Einstein gave a very simple interpretation of Lenard's results and borrowed Planck's hypothesis about the quantized energy from his blackbody research and assumed that the incoming radiation should be thought of as quanta of energy $h\nu$, with ν the frequency. In photoemission, one such quantum is absorbed by one electron. If the electron is some distance into the material of the cathode, some energy will be lost as it moves towards the surface. There will always be some electrostatic cost as the electron leaves the surface, which is the workfunction, Φ. The most energetic electrons emitted will be those very close to the surface, and they will leave the cathode with kinetic energy

$$KE = h\nu - \Phi \tag{1.3.1}$$

On cranking up the negative voltage on the collector plate until the current just stops, that is, to V_{stop}, the highest kinetic energy electrons (KE_e) must have had energy eV_{stop} upon leaving the cathode. Thus,

$$eV_{stop} = h\nu - \Phi \tag{1.3.2}$$

Thus, Einstein's theory makes a very definite quantitative prediction: if the frequency of the incident light is varied, and V_{stop} plotted as a function of frequency, the slope of the line should be $\frac{h}{e}$ (Figure **1.3.4*A***). It is also clear that there is a minimum light frequency for a given metal ν_o, that for which the quantum of energy is equal to Φ (Equation 1.3.1). Light below that frequency, no matter how bright, will not eject electrons.

> *According to both Planck and Einstein, the energy of light is proportional to its frequency rather than its amplitude, there will be a minimum frequency ν_0 needed to eject an electron with no residual energy.*

Since every photon of sufficient energy excites only one electron, increasing the light's intensity (i.e. the number of photons/sec) only increases the *number* of released electrons and not their kinetic energy. In addition, no time is necessary for the atom to be heated to a critical temperature and therefore the release of the electron is nearly instantaneous upon absorption of the light. Finally, because the photons must be above a certain energy to satisfy the workfunction, a threshold frequency exists below which no

photoelectrons are observed. This frequency is measured in units of Hertz (1/second) in honor of the discoverer of the photoelectric effect.

Einstein's Equation 1.3.1 explains the properties of the photoelectric effect quantitatively. A strange implication of this experiment is that light can behave as a kind of massless "particle" now known as a *photon* whose energy $E = h\nu$ can be transferred to an actual particle (an electron), imparting kinetic energy to it, just as in an elastic collision between to massive particles such as billiard balls.

> Robert Millikan initially did not accept Einstein's theory, which he saw as an attack on the wave theory of light, and worked for ten years until 1916, on the photoelectric effect. He even devised techniques for scraping clean the metal surfaces inside the vacuum tube. For all his efforts he found disappointing results: he confirmed Einstein's theory after ten years. In what he writes in his paper, Millikan is still desperately struggling to avoid this conclusion. However, by the time of his Nobel Prize acceptance speech, he has changed his mind rather drastically!

Einstein's simple explanation (Equation 1.3.1) completely accounted for the observed phenomena in Lenard's and Millikan's experiments (Figure 1.3.4) and began an investigation into the field we now call *quantum mechanics*. This new field seeks to provide a quantum explanation for classical mechanics and create a more unified theory of physics and thermodynamics. The study of the photoelectric effect has also lead to the creation of new field of photoelectron spectroscopy. Einstein's theory of the photoelectron presented a completely different way to measure Planck's constant than from black-body radiation.

The Workfunction (Φ)

The **workfunction** is an intrinsic property of the metal. While the workfunctions and ionization energies appear as similar concepts, they are independent. The workfunction of a metal is the minimum amount of energy (**E**) necessary to remove an electron from the surface of the bulk (*solid*) metal (sometimes referred to as **binding energy**).

$$\mathrm{M(s) + \Phi \rightarrow M^{+}(s) + e^{-}(\text{free with no kinetic energy})}$$

The workfunction is qualitatively similar to ionization energy (**IE**), which is the amount of energy required to remove an electron from an atom or molecule in the *gaseous* state.

$$\mathrm{M(g) + IE \rightarrow M^{+}(g) + e^{-}(\text{free with no kinetic energy})}$$

However, these two energies differ in magnitude (Table 1.3.1). For instance, copper has a workfunction of about 4.7 eV, but has a higher ionization energy of 7.7 eV. Generally, the ionization energies for metals are greater than the corresponding workfunctions (i.e., the electrons are less tightly bound in bulk metal).

Table 1.3.1 : Workfunctions and Ionization Energies of Select Elements

Element	Workfunction Φ (eV)	Ionization Energy (eV)
Lithium (Li)	2.93	5.39
Beryllium (Be)	4.98	9.32
Boron (B)	4.45	8.298
Carbon (C)	5.0	11.26
Sodium (Na)	2.36	5.13
Aluminum (Al)	4.20	5.98
Silicon (Si)	4.85	8.15
Potassium (K)	2.3	4.34
Iron (Fe)	4.67	7.87
Cobalt (Co)	4.89	7.88
Copper (Cu)	4.7	7.7
Gallium (Ga)	4.32	5.99
Germanium (Ge)	5.0	7.89
Arsenic (As)	3.75	9.81
Selenium (Se)	5.9	9.75
Silver (Ag)	4.72	7.57
Tin (Sn)	4.42	7.34
Cesium (Cs)	1.95	3.89
Gold (Au)	5.17	9.22
Mercury $(Hg)_{liquid}$	4.47	10.43
Bismuth (Bi)	4.34	7.29

Example 1.3.1 : Calcium

a. What is the energy in joules and electron volts of a photon of 420-nm violet light?
b. What is the maximum kinetic energy of electrons ejected from calcium by 420-nm violet light, given that the workfunction for calcium metal is 2.71 eV?

Strategy

To solve part (a), note that the energy of a photon is given by $E = h\nu$. For part (b), once the energy of the photon is calculated, it is a straightforward application of Equation 1.3.1 to find the ejected electron's maximum kinetic energy, since Φ is given.

Solution for (a)

Photon energy is given by

$$E = h\nu$$

Since we are given the wavelength rather than the frequency, we solve the familiar relationship $c = \nu\lambda$ for the frequency, yielding

$$\nu = \frac{c}{\lambda}$$

Combining these two equations gives the useful relationship

$$E = \frac{hc}{\lambda}$$

Now substituting known values yields

$$\begin{aligned} E &= \frac{(6.63 \times 10^{-34}\ J \cdot s)(3.00 \times 10^{8} m/s)}{420 \times 10^{-9}\ m} \\ &= 4.74 \times 10^{-19}\ J \end{aligned}$$

Converting to eV, the energy of the photon is

$$\begin{aligned} E &= (4.74 \times 10^{-19}\ J)\left(\frac{1\ eV}{1.6 \times 10^{-19}\ J}\right) \\ &= 2.96\ eV. \end{aligned}$$

Solution for (b)

Finding the kinetic energy of the ejected electron is now a simple application of Equation 1.3.1. Substituting the photon energy and binding energy yields

$$\begin{aligned} KE_e &= h\nu - \Phi \\ &= 2.96\ eV - 2.71\ eV \\ &= 0.246\ eV. \end{aligned}$$

Discussion

The energy of this 420-nm photon of violet light is a tiny fraction of a joule, and so it is no wonder that a single photon would be difficult for us to sense directly—humans are more attuned to energies on the order of joules. But looking at the energy in electron volts, we can see that this photon has enough energy to affect atoms and molecules. A DNA molecule can be broken with about 1 eV of energy, for example, and typical atomic and molecular energies are on the order of eV, so that the UV photon in this example could have biological effects.

The ejected electron (called a photoelectron) has a rather low energy, and it would not travel far, except in a vacuum. The electron would be stopped by a retarding potential of 0.26 eV. In fact, if the photon wavelength were longer and its energy less than 2.71 eV, then the formula would give a negative kinetic energy, an impossibility. This simply means that the 420-nm photons with their 2.96-eV energy are not much above the frequency threshold. You can show for yourself that the threshold wavelength is 459 nm (blue light). This means that if calcium metal is used in a light meter, the meter will be insensitive to wavelengths longer than those of blue light. Such a light meter would be insensitive to red light, for example.

? Exercise 1.3.1 : Silver

What is the longest-wavelength electromagnetic radiation that can eject a photoelectron from silver? Is this in the visible range?

? Exercise 1.3.2

Why is the workfunction of an element generally lower than the ionization energy of that element?

Summary

Although Hertz discovered the photoelectron in 1887, it was not until 1905 that a theory was proposed that explained the effect completely. The theory was proposed by Einstein and it made the claim that electromagnetic radiation had to be thought of as a series of particles, called photons, which collide with the electrons on the surface and emit them. This theory ran contrary to the belief that electromagnetic radiation was a wave and thus it was not recognized as correct until 1916 when Robert Millikan experimentally confirmed the theory

The photoelectric effect is the process in which electromagnetic radiation ejects electrons from a material. Einstein proposed photons to be quanta of electromagnetic radiation having energy $E = h\nu$ is the frequency of the radiation. All electromagnetic radiation is composed of photons. As Einstein explained, all characteristics of the photoelectric effect are due to the interaction of individual photons with individual electrons. The maximum kinetic energy KE_e of ejected electrons (photoelectrons) is given by $KE_e = h\nu - \Phi$, where $h\nu$ is the photon energy and Φ is the workfunction (or binding energy) of the electron to the particular material.

Conceptual Questions

1. Is visible light the only type of electromagnetic radiation that can cause the photoelectric effect?
2. Which aspects of the photoelectric effect cannot be explained without photons? Which can be explained without photons? Are the latter inconsistent with the existence of photons?
3. Is the photoelectric effect a direct consequence of the wave character of electromagnetic radiation or of the particle character of electromagnetic radiation? Explain briefly.
4. Insulators (nonmetals) have a higher Φ than metals, and it is more difficult for photons to eject electrons from insulators. Discuss how this relates to the free charges in metals that make them good conductors.
5. If you pick up and shake a piece of metal that has electrons in it free to move as a current, no electrons fall out. Yet if you heat the metal, electrons can be boiled off. Explain both of these facts as they relate to the amount and distribution of energy involved with shaking the object as compared with heating it.

Contributors and Attributions

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)
- Mark Tuckerman (New York University)
- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")

- Paul Flowers (University of North Carolina - Pembroke), Klaus Theopold (University of Delaware) and Richard Langley (Stephen F. Austin State University) with contributing authors. Textbook content produced by OpenStax College is licensed under a Creative Commons Attribution License 4.0 license. Download for free at http://cnx.org/contents/85abf193-2bd...a7ac8df6@9.110).

1.3: Photoelectric Effect Explained with Quantum Hypothesis is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.3: Photoelectric Effect Explained with Quantum Hypothesis** is licensed CC BY-NC-SA 4.0.

1.4: The Hydrogen Atomic Spectrum

Overview

- To introduce the concept of absorption and emission line spectra and describe the Balmer equation to describe the visible lines of atomic hydrogen.

The first person to realize that white light was made up of the colors of the rainbow was Isaac Newton, who in 1666 passed sunlight through a narrow slit, then a prism, to project the colored spectrum on to a wall. This effect had been noticed previously, of course, not least in the sky, but previous attempts to explain it, by Descartes and others, had suggested that the white light became colored when it was refracted, the color depending on the angle of refraction. Newton clarified the situation by using a second prism to reconstitute the white light, making much more plausible the idea that the white light was composed of the separate colors. He then took a monochromatic component from the spectrum generated by one prism and passed it through a second prism, establishing that no further colors were generated. That is, light of a single color did not change color on refraction. He concluded that white light was made up of all the colors of the rainbow, and that on passing through a prism, these different colors were refracted through slightly different angles, thus separating them into the observed spectrum.

Atomic Line Spectra

The spectrum of hydrogen *atoms*, which turned out to be crucial in providing the first insight into atomic structure over half a century later, was first observed by Anders Ångström in Uppsala, Sweden, in 1853. His communication was translated into English in 1855. Ångström, the son of a country minister, was a reserved person, not interested in the social life that centered around the court. Consequently, it was many years before his achievements were recognized, at home or abroad (most of his results were published in Swedish).

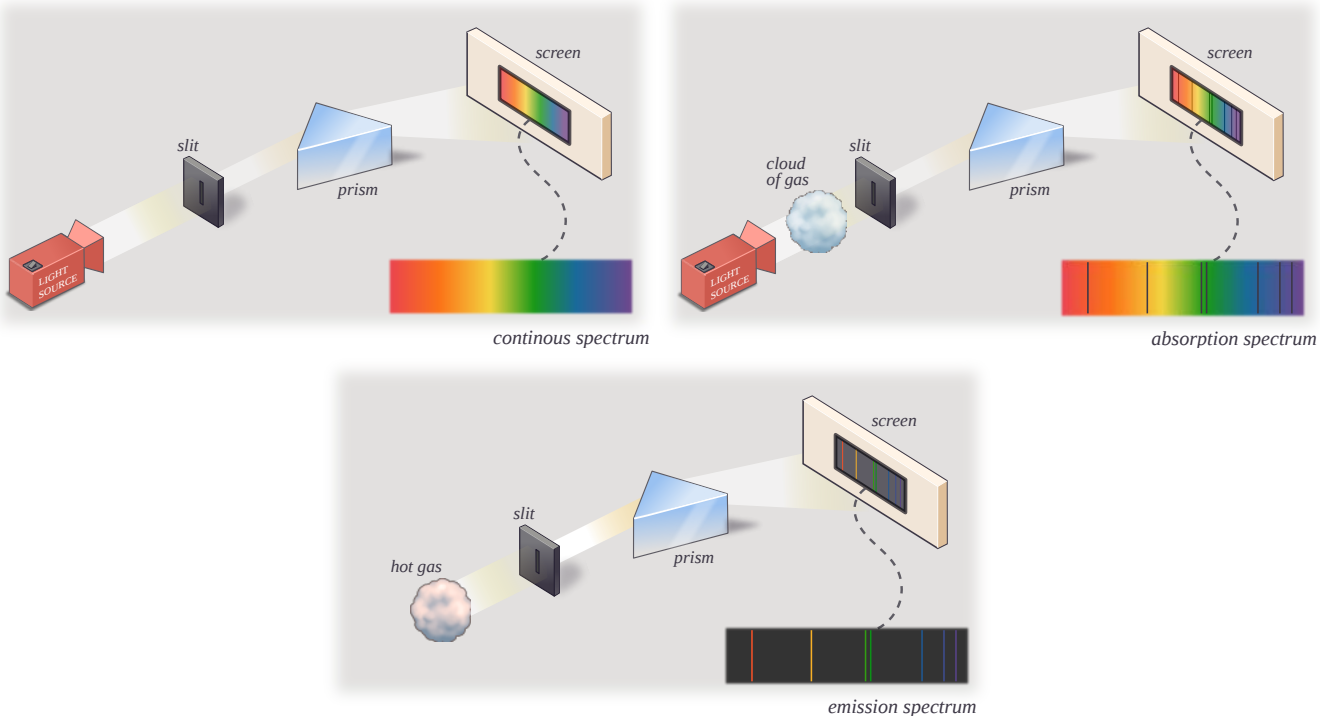


Figure 1.4.1 : Three Kinds of Spectra. When we see a lightbulb or other source of continuous radiation, all the colors are present. When the continuous spectrum is seen through a thinner gas cloud, the cloud's atoms produce absorption lines in the continuous spectrum. When the excited cloud is seen without the continuous source behind it, its atoms produce emission lines. We can learn which types of atoms are in the gas cloud from the pattern of absorption or emission lines. (CC BY-NC; Ümit Kaya via LibreTexts).

Most of what is known about atomic (and molecular) structure and mechanics has been deduced from spectroscopy. Figure 1.4.1 shows two different types of spectra. A continuous spectrum can be produced by an incandescent solid or gas at high pressure (e.g., blackbody radiation is a continuum). An emission spectrum can be produced by a gas at low pressure excited by heat or by collisions with electrons. An absorption spectrum results when light from a continuous source passes through a cooler gas, consisting of a series of dark lines characteristic of the composition of the gas.

Fraunhofer Lines

In 1802, William Wollaston in England had discovered that the solar spectrum had tiny gaps - there were many thin dark lines in the rainbow of colors. These were investigated much more systematically by Joseph von Fraunhofer, beginning in 1814. He increased the dispersion by using more than one prism. He found an "almost countless number" of lines. He labeled the strongest dark lines A, B, C, D, etc. Frauenhofer between 1814 and 1823 discovered nearly 600 dark lines in the solar spectrum viewed at high resolution and designated the principal features with the letters A through K, and weaker lines with other letters (Table 1.4.1). Modern observations of sunlight can detect many thousands of lines. It is now understood that these lines are caused by absorption by the outer layers of the Sun.

Table 1.4.1 : Major Fraunhofer lines and the elements they are associated with.

Designation	Element	Wavelength (nm)
y	O_2	898.765
Z	O_2	822.696
A	O_2	759.370
B	O_2	686.719
C	H	656.281
a	O_2	627.661
D_1	Na	589.592
D_2	Na	588.995
D_3 or d	He	587.5618

The Fraunhofer lines are typical spectral absorption lines. These dark lines are produced whenever a cold gas is between a broad spectrum photon source and the detector. In this case, a decrease in the intensity of light in the frequency of the incident photon is seen as the photons are absorbed, then re-emitted in random directions, which are mostly in directions different from the original one. This results in an *absorption line*, since the narrow frequency band of light initially traveling toward the detector, has been turned into heat or re-emitted in other directions.

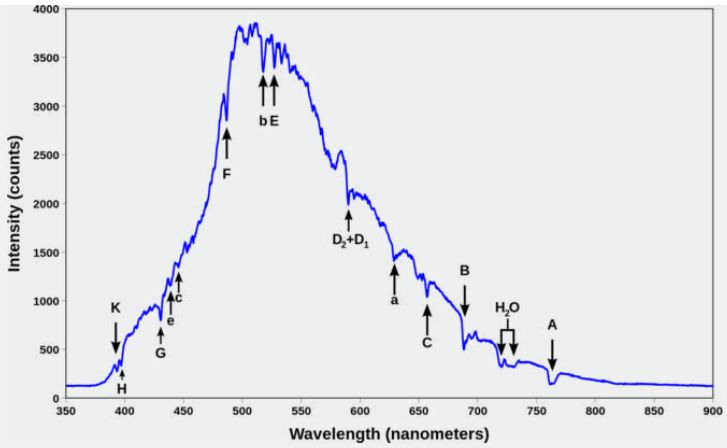


Figure 1.4.2 : Spectrum of blue sky. Dips are present at the Fraunhofer line wavelengths. (CC BY-SA 4.0; Eric Bajart via Wikipedia).

By contrast, if the detector sees photons emitted directly from a glowing gas, then the detector often sees photons emitted in a narrow frequency range by quantum emission processes in atoms in the hot gas, resulting in an *emission line*. In the Sun, Fraunhofer lines are seen from gas in the outer regions of the Sun, which are too cold to directly produce emission lines of the elements they represent.

Gases heated to incandescence were found by Bunsen, Kirkhoff and others to emit light with a series of sharp wavelengths. The emitted light analyzed by a spectrometer (or even a simple prism) appears as a multitude of narrow bands of color. These so called *line spectra* are characteristic of the atomic composition of the gas. The line spectra of several elements are shown in Figure 1.4.3 .

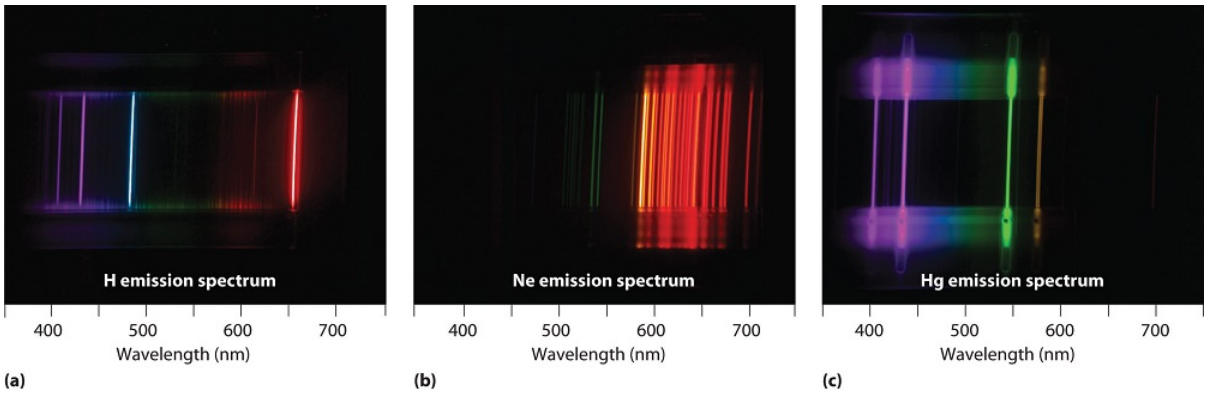


Figure 1.4.3 : The Emission Spectra of Elements Compared with Hydrogen. These images show (a) hydrogen gas, which is atomized to hydrogen atoms in the discharge tube; (b) neon; and (c) mercury. The strongest lines in the hydrogen spectrum are in the far UV Lyman series starting at 124 nm and below. The strongest lines in the mercury spectrum are at 181 and 254 nm, also in the UV; these are not shown. (CC BY-SA-NC; Anonymous by request).

The Balmer Series of Hydrogen

Obviously, if any pattern could be discerned in the spectral lines for a specific atom (in contract to the mixture that Fraunhofer lines represent), that might be a clue as to the internal structure of the atom. One might be able to build a model. A great deal of effort went into analyzing the spectral data from the 1860's on. The big breakthrough was made by Johann Balmer, a math and Latin teacher at a girls' school in Basel, Switzerland. Balmer had done no physics before and made his great discovery when he was almost sixty.

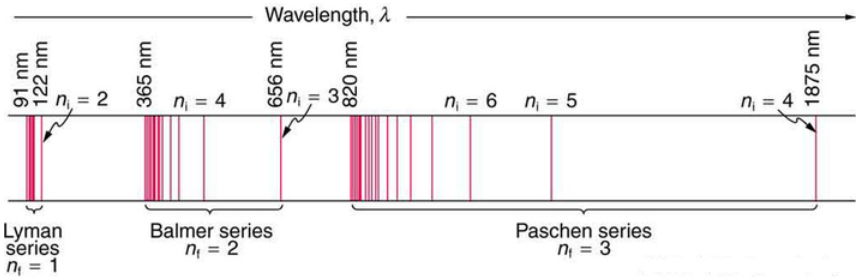


Figure 1.4.4 : A schematic of the hydrogen spectrum shows several series named for those who contributed most to their determination. Part of the Balmer series is in the visible spectrum, while the Lyman series is entirely in the UV, and the Paschen series and others are in the IR. Values of n_f and n_i are shown for some of the lines (CC BY-SA; OpenStax).

Balmer decided that the most likely atom to show simple spectral patterns was the lightest atom, hydrogen. Ångström had measured the four visible spectral lines to have wavelengths 656.21, 486.07, 434.01 and 410.12 nm (Figure 1.4.4). Balmer concentrated on just these four numbers, and found they were represented by the phenomenological formula:

$$\lambda = b\left(\frac{n_2^2}{n_2^2 - 4}\right) \tag{1.4.1}$$

where b = 364.56 nm and $n_2 = 3, 4, 5, 6$.

The first four wavelengths of Equation 1.4.1 (with n_2 = 3, 4, 5, 6) were in excellent agreement with the experimental lines from Ångström (Table 1.4.2). Balmer predicted that other lines exist in the ultraviolet that correspond to $n_2 \geq 7$ and in fact some of them had already been observed, unbeknown to Balmer.

Table 1.4.2 : The Balmer Series of Hydrogen Emission Lines

n_2	3	4	5	6	7	8	9	10
λ	*656*	*486*	*434*	*410*	397	389	383	380
color	*red*	*teal*	*blue*	*indigo*	violet	not visible	not visible	not visible

The n_2 integer in the Balmer series extends theoretically to infinity and the series represents a monotonically increasing energy (and frequency) of the absorption lines with increasing n_2 values. Moreover, the energy difference between successive lines decreased as n_2 increases (1.4.4). This behavior converges to a highest possible energy as Example 1.4.1 demonstrates. If the lines are plot according to their λ on a linear scale, you will get the appearance of the spectrum in Figure 1.4.4 ; these lines are called the **Balmer series**.

Balmer's general formula (Equation 1.4.1) can be rewritten in terms of the inverse wavelength typically called the *wavenumber* ($\tilde{\nu}$).

$$\tilde{\nu} = \frac{1}{\lambda} \tag{1.4.2}$$

$$= R_H\left(\frac{1}{4} - \frac{1}{n_2^2}\right) \tag{1.4.3}$$

where $n_2 = 3, 4, 5, 6$ and R_H is the Rydberg constant (discussed in the next section) equal to 109,737 cm^{-1}.

He further conjectured that the 4 could be replaced by 9, 16, 25, … and this also turned out to be true - but these lines, further into the infrared, were not detected until the early twentieth century, along with the ultraviolet lines.

The Wavenumber as a Unit of Frequency

The relation between wavelength and frequency for electromagnetic radiation is

$$\lambda\nu = c$$

In the SI system of units the wavelength, (λ) is measured in meters (m) and since wavelengths are usually very small one often uses the nanometer (nm) which is 10^{-9} m. The frequency (ν) in the SI system is measured in reciprocal seconds 1/s − which is

called a Hertz (after the discover of the photoelectron effect) and is represented by Hz.

It is common to use the reciprocal of the wavelength in centimeters as a measure of the frequency of radiation. This unit is called a wavenumber and is represented by ($\tilde{\nu}$) and is defined by

$$\begin{aligned}\tilde{\nu} &= \frac{1}{\lambda} \\ &= \frac{\nu}{c}\end{aligned}$$

Wavenumbers is a convenient unit in spectroscopy because it is directly proportional to energy.

$$\begin{aligned}E &= \frac{hc}{\lambda} \\ &= hc \times \frac{1}{\lambda} \\ &= hc\tilde{\nu} \\ &\propto \tilde{\nu}\end{aligned}$$

Example 1.4.1 : Balmer Series

Calculate the longest and shortest wavelengths (in nm) emitted in the Balmer series of the hydrogen atom emission spectrum.

Solution

From the behavior of the Balmer equation (Equation 1.4.1 and Table 1.4.2), the value of n_2 that gives the longest (i.e., greatest) wavelength (λ) is the smallest value possible of n_2, which is (n_2=3) for this series. This results in

$$\begin{aligned}\lambda_{longest} &= (364.56\ nm)\left(\frac{9}{9-4}\right) \\ &= (364.56\ nm)\,(1.8) \\ &= 656.2\ nm\end{aligned}$$

This is also known as the H_α line of atomic hydrogen and is bright red (Figure **1.4.3***a*).

For the shortest wavelength, it should be recognized that the shortest wavelength (greatest energy) is obtained at the limit of greatest (n_2):

$$\lambda_{shortest} = \lim_{n_2 \to \infty} (364.56\ nm)\left(\frac{n_2^2}{n_2^2-4}\right)$$

This can be solved via L'Hôpital's Rule, or alternatively the limit can be expressed via the equally useful energy expression (Equation 1.4.3) and simply solved:

$$\begin{aligned}\tilde{\nu}_{greatest} &= \lim_{n_2 \to \infty} R_H\left(\frac{1}{4} - \frac{1}{n_2^2}\right) \\ &= \lim_{n_2 \to \infty} R_H\left(\frac{1}{4}\right) \\ &= 27,434\ cm^{-1}\end{aligned}$$

Since $\frac{1}{\tilde{\nu}} = \lambda$ in units of cm, this converts to 364 nm as the shortest wavelength possible for the Balmer series.

The Balmer series is particularly useful in astronomy because the Balmer lines appear in numerous stellar objects due to the abundance of hydrogen in the universe, and therefore are commonly seen and relatively strong compared to lines from other elements.

1.4: The Hydrogen Atomic Spectrum is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.4: The Hydrogen Atomic Spectrum** by Michael Fowler is licensed CC BY-NC-SA 4.0.

1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum

Learning Objectives

- Describe Rydberg's theory for the hydrogen spectra.
- Interpret the hydrogen spectrum in terms of the energy states of electrons.

In an amazing demonstration of mathematical insight, in 1885 Balmer came up with a simple formula for predicting the wavelength of any of the lines in atomic hydrogen in what we now know as the Balmer series. Three years later, Rydberg generalized this so that it was possible to determine the wavelengths of any of the lines in the hydrogen emission spectrum. Rydberg suggested that all atomic spectra formed families with this pattern (he was unaware of Balmer's work). It turns out that there are families of spectra following Rydberg's pattern, notably in the alkali metals, sodium, potassium, etc., but not with the precision the hydrogen atom lines fit the Balmer formula, and low values of n_2 predicted wavelengths that deviate considerably.

Rydberg's phenomenological equation is as follows:

$$\tilde{\nu} = \frac{1}{\lambda} \tag{1.5.1}$$

$$= R_H \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \tag{1.5.2}$$

where R_H is the **Rydberg constant** and is equal to 109,737 cm^{-1} ($2.18 \times 10^{-18}\ J$) and n_1 and n_2 are integers (whole numbers) with $n_2 > n_1$.

For the Balmer lines, $n_1 = 2$ and n_2 can be any whole number between 3 and infinity. The various combinations of numbers that can be substituted into this formula allow the calculation the wavelength of any of the lines in the hydrogen emission spectrum; there is close agreement between the wavelengths generated by this formula and those observed in a real spectrum.

Other Series

The results given by Balmer and Rydberg for the spectrum in the visible region of the electromagnetic radiation start with $n_2 = 3$, and $n_1 = 2$. Is there a different series with the following formula (e.g., $n_1 = 1$)?

$$\frac{1}{\lambda} = R_H \left(\frac{1}{1^2} - \frac{1}{n^2} \right) \tag{1.5.3}$$

The values for n_2 and wavenumber $\tilde{\nu}$ for this series would be:

Table 1.5.1 : The Lyman Series of Hydrogen Emission Lines ($n_1 = 1$)

n_2	2	3	4	5	...
λ (nm)	121	102	97	94	...
$\tilde{\nu}$ (cm^{-1})	82,2291	97,530	102,864	105,332	...

Do you know in what region of the electromagnetic radiation these lines are? Of course, these lines are in the UV region, and they are not visible, but they are detected by instruments; these lines form a **Lyman series**. The existences of the Lyman series and Balmer's series suggest the existence of more series. For example, the series with $n_1 = 3$ and $n_2 = 4, 5, 6, 7, ...$ is called **Paschen series**.

Other Series

The spectral lines are grouped into series according to n_1 values. Lines are named sequentially starting from the longest wavelength/lowest frequency of the series, using Greek letters within each series. For example, the ($n_1 = 1/n_2 = 2$) line is called "Lyman-alpha" (Ly-α), while the ($n_1 = 3/n_2 = 7$) line is called "Paschen-delta" (Pa-δ). The first six series have specific names:

- Lyman series with $n_1 = 1$

- Lyman series with $n_1 = 1$
- Balmer series with $n_1 = 2$
- Paschen series (or Bohr series) with $n_1 = 3$
- Brackett series with $n_1 = 4$
- Pfund series with $n_1 = 5$
- Humphreys series with $n_1 = 6$

Example 1.5.1 : The Lyman Series

The so-called Lyman series of lines in the emission spectrum of hydrogen corresponds to transitions from various excited states to the n = 1 orbit. Calculate the wavelength of the lowest-energy line in the Lyman series to three significant figures. In what region of the electromagnetic spectrum does it occur?

Given: lowest-energy orbit in the Lyman series

Asked for: wavelength of the lowest-energy Lyman line and corresponding region of the spectrum

Strategy

A. Substitute the appropriate values into Equation 1.5.2 (the Rydberg equation) and solve for λ.

B. Locate the region of the electromagnetic spectrum corresponding to the calculated wavelength.

Solution

We can use the Rydberg equation (Equation 1.5.2) to calculate the wavelength:

$$\frac{1}{\lambda} = R_H \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right)$$

Solution A

For the Lyman series, $n_1 = 1$.

$$\begin{aligned} \frac{1}{\lambda} &= R_H \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \\ &= 1.097 \times 10^7 \, m^{-1} \left(\frac{1}{1} - \frac{1}{4} \right) \\ &= 8.228 \times 10^6 \, m^{-1} \end{aligned}$$

Spectroscopists often talk about energy and frequency as equivalent. The cm^{-1} unit (wavenumbers) is particularly convenient. We can convert the answer in part A to cm^{-1}.

$$\begin{aligned} \tilde{\nu} &= \frac{1}{\lambda} \\ &= 8.228 \times 10^6 \, \cancel{m^{-1}} \left(\frac{\cancel{m}}{100 \, cm} \right) \\ &= 82,280 \, cm^{-1} \end{aligned}$$

and

$$\lambda = 1.215 \times 10^{-7} \, m = 122 \, nm$$

This emission line is called Lyman alpha and is the strongest atomic emission line from the sun and drives the chemistry of the upper atmosphere of all the planets producing ions by stripping electrons from atoms and molecules. It is completely absorbed by oxygen in the upper stratosphere, dissociating O_2 molecules to O atoms which react with other O_2 molecules to form stratospheric ozone

Solution B

his wavelength is in the ultraviolet region of the spectrum.

Exercise 1.5.1 : The Pfund Series

The Pfund series of lines in the emission spectrum of hydrogen corresponds to transitions from higher excited states to the $n_1 = 5$. Calculate the wavelength of the *second* line in the Pfund series to three significant figures. In which region of the spectrum does it lie?

The above discussion presents only a phenomenological description of hydrogen emission lines and fails to provide a probe of the nature of the atom itself. Clearly a continuum model based on classical mechanics is not applicable, and as the next Section demonstrates, a simple connection between spectra and atomic structure can be formulated.

1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum** by Michael Fowler is licensed CC BY-NC-SA 4.0.

1.6: Matter Has Wavelike Properties

Learning Objectives

- To introduce the wave-particle duality of light extends to matter
- To describe how matter (e.g., electrons and protons) can exhibit wavelike properties, e.g., interference and diffraction patterns
- To use algebra to find the de Broglie wavelength or momentum of a particle when either one of these quantities is given

The next real advance in understanding the atom came from an unlikely quarter - a student prince in Paris. Prince Louis de Broglie was a member of an illustrious family, prominent in politics and the military since the 1600's. Louis began his university studies with history, but his elder brother Maurice studied x-rays in his own laboratory, and Louis became interested in physics. After World War I, de Broglie focused his attention on Einstein's two major achievements, the theory of special relativity and the quantization of light waves. He wondered if there could be some connection between them. Perhaps the quantum of radiation really should be thought of as a particle. De Broglie suggested that if waves (photons) could behave as particles, as demonstrated by the photoelectric effect, then the converse, namely that particles could behave as waves, should be true. He associated a wavelength λ to a particle with momentum p using Planck's constant as the constant of proportionality:

$$\lambda = \frac{h}{p} \tag{1.6.1}$$

which is called the **de Broglie wavelength**. The fact that particles can behave as waves but also as particles, depending on which experiment you perform on them, is known as the **wave-particle duality**.

Deriving the de Broglie Wavelength

From the discussion of the photoelectric effect, we have the first part of the particle-wave duality, namely, that electromagnetic waves can behave like particles. These particles are known as *photons*, and they move at the speed of light. Any particle that moves at or near the speed of light has kinetic energy given by Einstein's special theory of relatively. In general, a particle of mass m and momentum p has an energy

$$E = \sqrt{p^2c^2 + m^2c^4} \tag{1.6.2}$$

Note that if $p = 0$, this reduces to the famous rest-energy expression $E = mc^2$. However, photons are massless particles (technically rest-massless) that always have a finite momentum p. In this case, Equation 1.6.2 becomes

$$E = pc.$$

From Planck's hypothesis, one quantum of electromagnetic radiation has energy $E = h\nu$. Thus, equating these two expressions for the kinetic energy of a photon, we have

$$h\nu = \frac{hc}{\lambda} = pc \tag{1.6.3}$$

Solving for the wavelength λ gives Equation 1.6.1:

$$\lambda = \frac{h}{p} = \frac{h}{mv}$$

where v is the velocity of the particle. Hence, de Broglie argued that if particles can behave as waves, then a relationship like this, which pertains particularly to waves, should also apply to particles.

Equation 1.6.1 allows us to associate a wavelength λ to a particle with momentum p. As the momentum increases, the wavelength decreases. In both cases, this means the energy becomes larger. i.e., short wavelengths *and* high momenta correspond to high energies.

> *It is a common feature of quantum mechanics that particles and waves with short wavelengths correspond to high energies and vice versa.*

Having decided that the photon might well be a particle with a rest mass, even if very small, it dawned on de Broglie that in other respects it might not be too different from other particles, especially the very light electron. In particular, maybe the electron also had an associated wave. The obvious objection was that if the electron was wavelike, why had no diffraction or interference effects been observed? But there was an answer. If de Broglie's relation between momentum and wavelength also held for electrons, the wavelength was sufficiently short that these effects would be easy to miss. As de Broglie himself pointed out, the wave nature of light is not very evident in everyday life. As the next section will demonstrate, the validity of de Broglie's proposal was confirmed by electron diffraction experiments of G.P. Thomson in 1926 and of C. Davisson and L. H. Germer in 1927. In these experiments it was found that electrons were scattered from atoms in a crystal and that these scattered electrons produced an interference pattern. These diffraction patterns are characteristic of wave-like behavior and are exhibited by both electrons (i.e., matter) and electromagnetic radiation (i.e., light).

Example 1.6.1 : Electron Waves

Calculate the de Broglie wavelength for an electron with a kinetic energy of 1000 eV.

Solution

To calculate the de Broglie wavelength (Equation 1.6.1), the momentum of the particle must be established and requires knowledge of both the mass and velocity of the particle. The mass of an electron is $9.109383 \times 10^{-28}\ g$ and the velocity is obtained from the given kinetic energy of 1000 eV:

$$\begin{aligned} KE &= \frac{mv^2}{2} \\ &= \frac{p^2}{2m} = 1000\ eV \end{aligned}$$

Solve for momentum

$$p = \sqrt{2mKE}$$

convert to SI units

$$p = \sqrt{(1000\ \cancel{eV})\left(\frac{1.6 \times 10^{-19}\ J}{1\ \cancel{eV}}\right)(2)(9.109383 \times 10^{-31}\ kg)}$$

expanding definition of joule into base SI units and cancel

$$\begin{aligned} p &= \sqrt{(3.1 \times 10^{-16}\ kg \cdot m^2/s^2)(9.109383 \times 10^{-31}\ kg)} \\ &= \sqrt{2.9 \times 10^{-40}\ kg^2\ m^2/s^2} \\ &= 1.7 \times 10^{-23} kg \cdot m/s \end{aligned}$$

Now substitute the momentum into the equation for de Broglie's wavelength (Equation 1.6.1) with Planck's constant ($h = 6.626069 \times 10^{-34}\ J \cdot s$). After expanding units in Plank's constant

$$\begin{aligned} \lambda &= \frac{h}{p} \\ &= \frac{6.626069 \times 10^{-34}\ kg \cdot m^2/s}{1.7 \times 10^{-23} kg \cdot m/s} \\ &= 3.87 \times 10^{-11}\ m \\ &= 38.9\ pm \end{aligned}$$

Exercise 1.6.1 : Baseball Waves

Calculate the de Broglie wavelength for a fast ball thrown at 100 miles per hour and weighing 4 ounces. Comment on whether the wave properties of baseballs could be experimentally observed.

Exercise 1.6.2 : Electrons vs. Protons

If an electron and a proton have the same velocity, which would have the longer de Broglie wavelength?

a. The electron
b. The proton
c. They would have the same wavelength

This was the prince's Ph.D. thesis, presented in 1924. His thesis advisor was somewhat taken aback, and was not sure if this was sound work. He asked de Broglie for an extra copy of the thesis, which he sent to Einstein. Einstein wrote shortly afterwards: *"I believe it is a first feeble ray of light on this worst of our physics enigmas*" and the prince got his Ph.D.

1.6: Matter Has Wavelike Properties is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.6: Matter Has Wavelike Properties** by Michael Fowler is licensed CC BY-NC-SA 4.0.

1.7: de Broglie Waves can be Experimentally Observed

Learning Objectives

- To present the experimental evidence behind the wave-particle duality of matter

The validity of de Broglie's proposal was confirmed by electron diffraction experiments of G.P. Thomson in 1926 and of C. Davisson and L. H. Germer in 1927. In these experiments it was found that electrons were scattered from atoms in a crystal and that these scattered electrons produced an interference pattern. The interference pattern was just like that produced when water waves pass through two holes in a barrier to generate separate wave fronts that combine and interfere with each other. These diffraction patterns are characteristic of wave-like behavior and are exhibited by both matter (e.g., electrons and neutrons) and electromagnetic radiation. Diffraction patterns are obtained if the wavelength is comparable to the spacing between scattering centers.

> *Diffraction occurs when waves encounter obstacles whose size is **comparable** with its wavelength.*

Continuing with our analysis of experiments that lead to the new quantum theory, we now look at the phenomenon of electron diffraction.

Diffraction of Light (Light as a Wave)

It is well-known that *light* has the ability to diffract around objects in its path, leading to an interference pattern that is particular to the object. This is, in fact, how holography works (the interference pattern is created by allowing the diffracted light to interfere with the original beam so that the hologram can be viewed by shining the original beam on the image). A simple illustration of light diffraction is the Young double slit experiment (Figure 1.7.1). Here, light as waves (pictured as waves in a plane parallel to the double slit apparatus) impinge on the two slits. Each slit then becomes a point source for spherical waves that subsequently interfere with each other, giving rise to the light and dark fringes on the screen at the right.

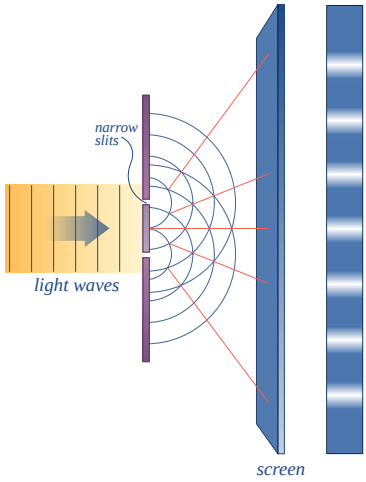


Figure 1.7.1 : Young double slit experiment. Two slits are illuminated by a plane light waves. (CC BY-NC; Ümit Kaya via LibreTexts)

The double-slit experiments are direct demonstration of wave phenomena via observed interference. These types of experiment were first performed by Thomas Young in 1801, as a demonstration of the wave behavior of light. In the basic version of this experiment, light is illuminated only a plate pierced by two parallel slits, and the light passing through the two slits is observed on a screen behind the plate (Figures 1.7.1 and 1.7.2).

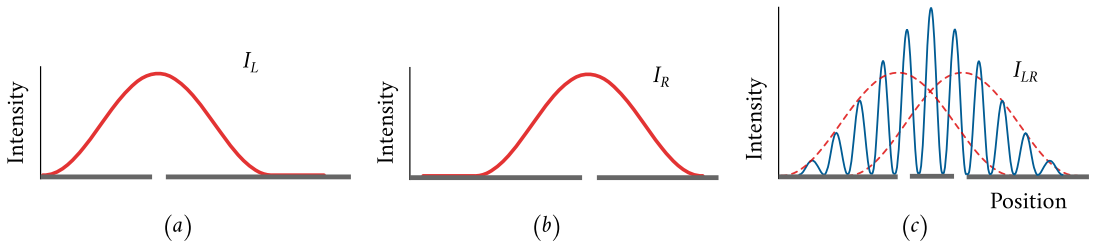


Figure 1.7.2 : Amplitude of particles through (a) left slit, (b) right slit, (c) through both slits. In absence of wavelength interference behavior, the intensity will be the sum of each single slit results (red) and when there is interference, pronounced oscillations in intensity is observed (blue curve). (CC BY; Ümit Kaya via LibreTexts)

The wave nature of light causes the light waves passing through the two slits to interfere, producing bright and dark bands on the screen – a result that would not be expected if light consisted of classical particles (Figure **1.7.0c**). However, the light is always found to be absorbed at the screen at discrete points, as individual particles (not waves); the interference pattern appears via the varying density of these particle hits on the screen. Furthermore, versions of the experiment that include detectors at the slits find that each detected photon passes through one slit (as would a classical particle), and not through both slits (as would a wave).

> *Interference is a wave phenomenon in which two (or more) waves superimpose to form a resultant wave of greater or lower amplitude. It is the primary property used to identify wave behavior.*

Diffraction of Electrons (Electrons as Waves)

According to classical physics, electrons should behave like particles - they travel in straight lines and do not curve in flight unless acted on by an external agent, like a magnetic field. In this model, if we fire a beam of electrons through a double slit onto a detector, we should get two bands of "hits", much as you would get if you fired a machine gun at the side of a house with two windows - you would get two areas of bullet-marked wall inside, and the rest would be intact (Figure 1.7.3 (left) and Figure 1.7.2).

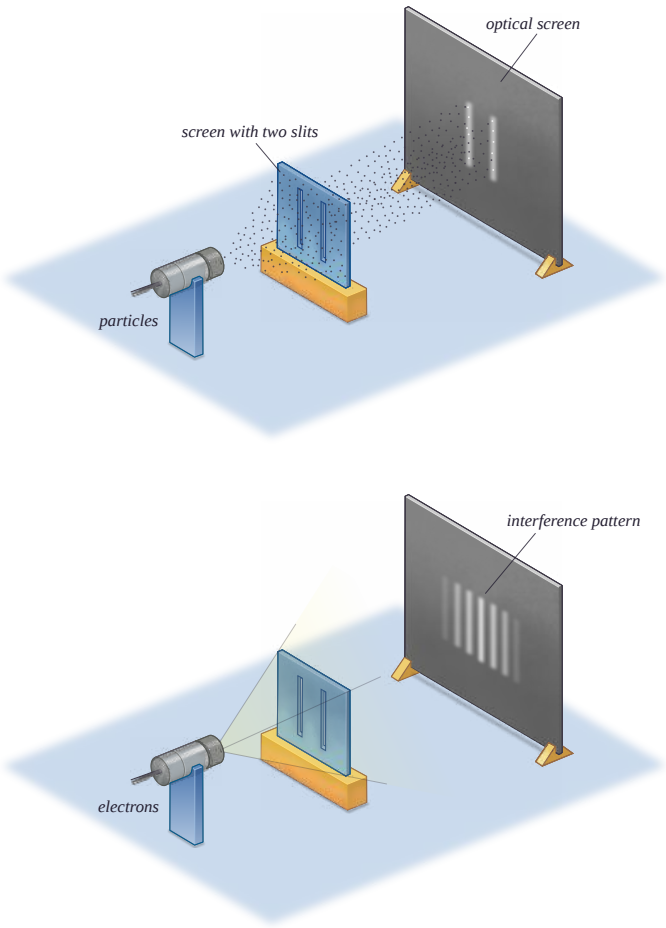


Figure 1.7.3 : (left) classical model of electrons. (right) wave property of electrons. (CC BY-NC; Ümit Kaya via LibreTexts)

However, if the slits are made small enough and close enough together, we actually observe the electrons are *diffracting* through the slits and *interfering* with each other just like waves (Figure 1.7.3 (right) and Figure 1.7.2 a,b). This means that the electrons have **wave-particle duality**, just like photons, in agreement with de Broglie's hypothesis discussed previously. In this case, they must have properties like wavelength and frequency. We can deduce the properties from the behavior of the electrons as they pass through our diffraction grating.

This was a pivotal result in the development of quantum mechanics. Just as the photoelectric effect demonstrated the particle nature of light, the Davisson–Germer experiment showed the wave-nature of matter, and completed the theory of wave-particle duality. For physicists this idea was important because it meant that not only could any particle exhibit wave characteristics, but that one could use wave equations to describe phenomena in matter if one used the de Broglie wavelength.



Figure 1.7.4 : An image of an ant in a scanning electron microscope based on the wave properties of electrons. (Public Domain; United States Geological Survey, an agency of the United States Department of the Interior)

An electron microscope use a beam of accelerated electrons as a source of illumination. Since the wavelength of electrons can be up to 100,000 times shorter than that of visible light photons, electron microscopes have a higher resolving power than light microscopes and can reveal the structure of smaller objects. A transmission electron microscope can achieve better than 50 pm resolution and magnifications of up to about 10,000,000x whereas most light microscopes are limited by diffraction to about 200 nm resolution and useful magnifications below 2000x (Figure 1.7.4).

Is Matter a Particle or a Wave?

An electron, indeed any particle, is neither a *particle* nor a *wave*. Describing the electron as a particle is a mathematical model that works well in some circumstances while describing it as a wave is a different mathematical model that works well in other circumstances. When you choose to do some calculation of the electron's behavior that treats it either as a particle or as a wave, you're not saying the electron **is** a particle or **is** a wave: you're just choosing the mathematical model that makes it easiest to do the calculation.

Neutron Diffraction (Neutrons as Waves)

Like all quantum particles, neutrons can also exhibit wave phenomena and if that wavelength is short enough, atoms or their nuclei can serve as diffraction obstacles. When a beam of neutrons emanating from a reactor is slowed down and selected properly by their speed, their wavelength lies near one angstrom (0.1 nanometer), the typical separation between atoms in a solid material. Such a beam can then be used to perform a diffraction experiment. Neutrons interact directly with the nucleus of the atom, and the contribution to the diffracted intensity depends on each isotope; for example, regular hydrogen and deuterium contribute differently. It is also often the case that light (low Z) atoms contribute strongly to the diffracted intensity even in the presence of large Z atoms.

Example 1.7.1 : Neutron Diffraction

Neutrons have no electric charge, so they do not interact with the atomic electrons. Hence, they are very penetrating (e.g., typically 10 cm in lead). Neutron diffraction was proposed in 1934, to exploit de Broglie's hypothesis about the wave nature of matter. Calculate the momentum and kinetic energy of a neutron whose wavelength is comparable to atomic spacing ($1.8 \times 10^{-10}\, m$).

Solution

This is a simple use of de Broglie's equation

$$\lambda = \frac{h}{p}$$

where we recognize that the wavelength of the neutron must be comparable to atomic spacing (let's assumed equal for convenience, so $\lambda = 1.8 \times 10^{-10}\, m$). Rearranging the de Broglie wavelength relationship above to solve for momentum (p):

$$\begin{aligned} p &= \frac{h}{\lambda} \\ &= \frac{6.6 \times 10^{-34} Js}{1.8 \times 10^{-10} m} \\ &= 3.7 \times 10^{-24}\, kg\ m\ s^{-1} \end{aligned}$$

The relationship for kinetic energy is

$$KE = \frac{1}{2}mv^2 = \frac{p^2}{2m}$$

where v is the velocity of the particle. From the reference table of physical constants, the mass of a neutron is $1.6749273 \times 10^{-27}$ kg, so

$$KE = \frac{(3.7 \times 10^{-24}\ kg\ m\ s^{-1})^2}{2(1.6749273 \times 10^{-27}\ kg)}$$
$$= 4.0 \times 10^{-21} J$$

The neutrons released in nuclear fission are ‘fast’ neutrons, i.e. much more energetic than this. Their wavelengths be much smaller than atomic dimensions and will not be useful for neutron diffraction. We slow down these fast neutrons by introducing a "moderator", which is a material (e.g., graphite) that neutrons can penetrate, but will slow down appreciably.

1.7: de Broglie Waves can be Experimentally Observed is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.7: de Broglie Waves can be Experimentally Observed** by John Rennie, Mark Tuckerman, Michael Fowler is licensed CC BY-NC-SA 4.0.

1.8: The Bohr Theory of the Hydrogen Atom

Learning Objectives

- Introduce the fundamentals behind the Bohr Atom and demonstrate it can predict the Rydberg's equation for the atomic spectrum of hydrogen

Rutherford's Failed Planetary Atom

Ernest Rutherford had proposed a model of atoms based on the α-particle scattering experiments of Hans Geiger and Ernest Marsden. In these experiments helium nuclei (α-particles) were shot at thin gold metal foils. Most of the particles were not scattered; they passed unchanged through the thin metal foil. Some of the few that were scattered were scattered in the backward direction; i.e. they recoiled. This backward scattering requires that the foil contain heavy particles. When an α-particle hits one of these heavy particles it simply recoils backward, just like a ball thrown at a brick wall. Since most of the α-particles don't get scattered, the heavy particles (the nuclei of the atoms) must occupy only a very small region of the total space of the atom. Most of the space must be empty or occupied by very low-mass particles. These low-mass particles are the electrons that surround the nucleus.

There are some basic problems with the Rutherford model. The Coulomb force that exists between oppositely charge particles means that a positive nucleus and negative electrons should attract each other, and the atom should collapse. To prevent the collapse, the electron was postulated to be orbiting the positive nucleus. The Coulomb force (discussed below) is used to change the direction of the velocity, just as a string pulls a ball in a circular orbit around your head or the gravitational force holds the moon in orbit around the Earth. The origin for this hypothesis that suggests this perspective is plausible is the similarity of gravity and Coulombic interactions. The expression for the force of gravity between two masses (Newton's Law of gravity) is

$$F_{gravity} \propto \frac{m_1 m_2}{r^2} \tag{1.8.1}$$

with m_1 and m_2 representing the **mass** of object 1 and 2, respectively and r representing the distance between the objects centers

The expression for the Coulomb force between two charged species is

$$F_{Coulomb} \propto \frac{Q_1 Q_2}{r^2} \tag{1.8.2}$$

with Q_1 and Q_2 representing the **charge** of object 1 and 2, respectively and r representing the distance between the objects centers.

However, this analogy has a problem too. An electron going around in a circle is constantly being accelerated because its velocity vector is changing. A charged particle that is being accelerated emits radiation. This property is essentially how a radio transmitter works. A power supply drives electrons up and down a wire and thus transmits energy (electromagnetic radiation) that your radio receiver picks up. The radio then plays the music for you that is encoded in the waveform of the radiated energy.

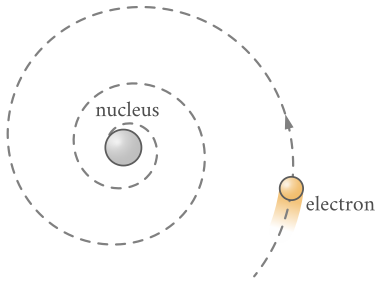


Figure 1.8.1 : The classical death spiral of an electron around a nucleus. (CC BY-NC; Ümit Kaya via LibreTexts)

If the orbiting electron is generating radiation, it is losing energy. If an orbiting particle loses energy, the radius of the orbit decreases. To conserve angular momentum, the frequency of the orbiting electron increases. The frequency increases continuously as the electron collapses toward the nucleus. Since the frequency of the rotating electron and the frequency of the radiation that is

emitted are the same, both change continuously to produce a continuous spectrum and not the observed discrete lines. Furthermore, if one calculates how long it takes for this collapse to occur, one finds that it takes about 10^{-11} seconds. This means that nothing in the world based on the structure of atoms could exist for longer than about 10^{-11} seconds. Clearly something is terribly wrong with this classical picture, which means that something was missing at that time from the known laws of physics.

Conservative Forces can be explained with Potentials

A conservative force is dependent only on the position of the object. If a force is conservative, it is possible to assign a numerical value for the potential at any point. When an object moves from one location to another, the force changes the potential energy of the object by an amount that does not depend on the path taken. The potential can be constructed as simple derivatives for 1-D forces:

$$F = -\frac{dV}{dx}$$

or as gradients in 3-D forces

$$F = -\nabla V$$

where ∇ is the vector of partial derivatives

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

The most familiar conservative forces are gravity and Coulombic forces.

The Coulomb force law (Equation 1.8.2) comes from the corresponding Coulomb potential (sometimes call electrostatic potential)

$$V(r) = \frac{kQ_1Q_2}{r} \tag{1.8.3}$$

and it can be easily verified that the Coulombic force from this interaction ($F(r)$) is

$$F(r) = -\frac{dV}{dr} \tag{1.8.4}$$

As r is varied, the energy will change, so that we have an example of a potential energy curve $V(r)$ (Figure 1.8.2;). If Q_1 and Q_2 are the same sign, then the curve which is a purely **repulsive potential**, i.e., the energy increases monotonically as the charges are brought together and decreases monotonically as they are separated. From this, it is easy to see that like charges (charges of the same sign) repel each other.

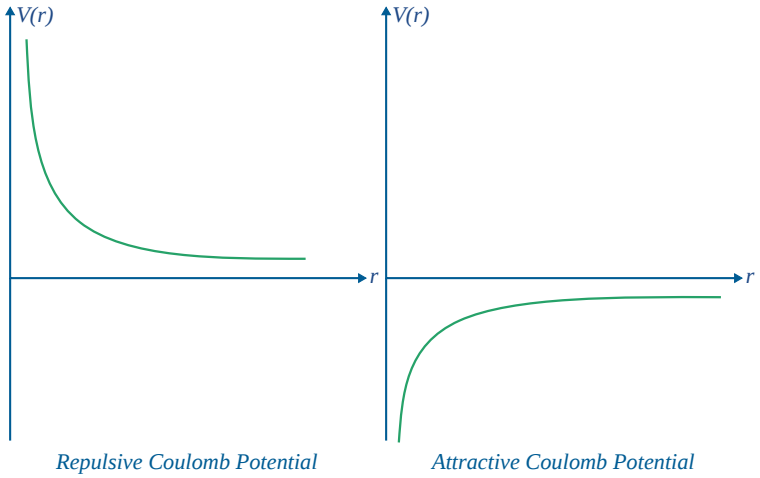


Figure 1.8.2 : Potential energy curve for the Coulomb interactions between two charges of same sign (left) and opposite signs (right). (CC BY-NC; Ümit Kaya via LibreTexts)

If the charges are of opposite sign, then the curve appears roughly Figure **1.8.2;** and this is a purely **attractive potential**. Thus, the energy decreases as the charges are brought together, implying that opposite charges attract

The Bohr Model

It is observed that line spectra discussed in the previous sections show that hydrogen atoms absorb and emit light at only discrete wavelengths. This observation is connected to the discrete nature of the allowed energies of a quantum mechanical system. Quantum mechanics postulates that, in contrast to classical mechanics, the energy of a system can only take on certain discrete values. This leaves us with the question: How do we determine what these allowed discrete energy values are? After all, it seems that Planck's formula for the allowed energies came out of nowhere.

The model we will describe here, due to Niels Bohr in 1913, is an early attempt to predict the allowed energies for single-electron atoms such as **H**, $\mathbf{He^{+}}$, $\mathbf{Li^{2+}}$, $\mathbf{Be^{3+}}$, etc. Although Bohr's reasoning relies on classical concepts and hence, is not a correct explanation, the reasoning is interesting, and so we examine this model for its historical significance.

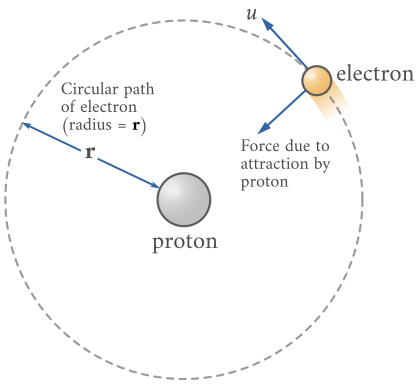


Figure 1.8.3 : Bohr atom with an electron revolving around a fixed nucleus. (CC BY-NC; Ümit Kaya via LibreTexts)

Consider a nucleus with charge $+Ze$ and one electron orbiting the nucleus. In this analysis, we will use another representation of the constant k in Coulomb's law (Equation 1.8.3), which is more commonly represented in the form:

$$k = \frac{1}{4\pi\epsilon_0} \tag{1.8.5}$$

where ϵ_0 is known as the **permittivity of free space** with the numerical value $\epsilon_0 = 8.8541878 \times 10^{-12}\ C^2 J^{-1} m^{-1}$.

The total energy of the electron (the nucleus is assumed to be fixed in space at the origin) is the sum of kinetic and potential energies:

$$E_{total} = \underbrace{\frac{p^2}{2m_e}}_{\text{kinetic energy}} - \underbrace{\frac{Ze^2}{4\pi\epsilon_0 r}}_{\text{potential energy}}$$

The force on the electron is

$$\vec{F} = -\frac{Ze^2}{4\pi\epsilon_0 r^3} r$$

and its magnitude is

$$F = |\vec{F}| = \frac{Ze^2}{4\pi\epsilon_0 r^3}|r| = \frac{Ze^2}{4\pi\epsilon_0 r^2}$$

since $\vec{F} = m_e\vec{a}$, the magnitude, it follows that $|\vec{F}| = m_e|\vec{a}|$. If we assume that the orbit is circular, then the acceleration is *purely centripetal*, so

$$|a| = \frac{v^2}{r}$$

where v is the velocity of the electron. Equating force $|F|$ to $m_e|a|$, we obtain

$$\frac{Ze^2}{4\pi\epsilon_0 r^2} = m_e\frac{v^2}{r}$$

or

$$\frac{Ze^2}{4\pi\epsilon_0} = m_e v^2 r$$

or

$$\frac{Ze^2 m_e r}{4\pi\epsilon_0} = (m_e v r)^2 \tag{1.8.6}$$

The reason for writing the equation this way is that the quantity $m_e vr$ is the classical orbital angular momentum of the electron. Bohr was familiar with Maxwell's theory of classical electromagnetism and knew that in a classical theory, the orbiting electron should radiate energy away and eventually collapse into the nucleus (Figure 1.8.1). He circumvented this problem by following Planck's idea underlying blackbody radiation and positing that the orbital angular momentum $m_e vr$ of the electron could only take on specific values

$$m_e vr = n\hbar \tag{1.8.7}$$

with $n = 1, 2, 3, ...$

> *Note that the electron must be in motion, so $n = 0$ is not allowed.*

Substituting Equation 1.8.7 into the Equation 1.8.6, we find

$$\frac{Ze^2 m_e r}{4\pi\epsilon_0} = n^2(\hbar)^2 \tag{1.8.8}$$

Equation 1.8.8 implies that orbits could only have certain allowed radii

$$r_n = \frac{4\pi\epsilon_0\hbar^2}{Ze^2 m_e}n^2 \tag{1.8.9}$$

$$= \frac{a_0}{Z}n^2 \tag{1.8.10}$$

with $n = 1, 2, 3, ...$ The collection of constants has been defined to be a_0

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{e^2 m_e}$$

a quantity that is known as the **Bohr radius**.

We can also calculate the allowed momenta since $m_e vr = n\hbar$, and $p = m_e v$. Thus,

$$p_n r_n = n\hbar$$

$$p_n = \frac{n\hbar}{r_n}$$

$$= \frac{\hbar Z}{a_0 n}$$

$$= \frac{Ze^2 m_e}{4\pi\epsilon_0\hbar n}$$

From p_n and r_n, we can calculate the allowed energies from

$$E_n = \frac{p_n^2}{2m_e} - \frac{Ze^2}{4\pi\epsilon_0 r_n}$$

Substituting in the expressions for p_n and r_n and simplifying gives

$$E_n = -\frac{Z^2 e^4 m_e}{32\pi^2 \epsilon_0^2 \hbar^2}\frac{1}{n^2} = -\frac{e^4 m_e}{8\epsilon_0^2 h^2}\frac{Z^2}{n^2} \tag{1.8.11}$$

We can redefine a new energy scale by defining **the Rydberg** as

$$1\ Ry = \frac{e^4 m_e}{8\epsilon_0^2 h^2} = 2.18 \times 10^{-18}\ J.$$

and this simplifies the allowed energies predicted by the Bohr model (Equation 1.8.11) as

$$E_n = -(2.18 \times 10^{-18})\frac{Z^2}{n^2}\ J = -\frac{Z^2}{n^2}\ R_y \tag{1.8.12}$$

Hence, the energy of the electron in an atom also is quantized. Equation 1.8.12 gives the energies of the electronic states of the hydrogen atom. It is very useful in analyzing spectra to represent these energies graphically in an energy-level diagram. An energy-level diagram has energy plotted on the vertical axis with a horizontal line drawn to locate each energy level (Figure 1.8.4).

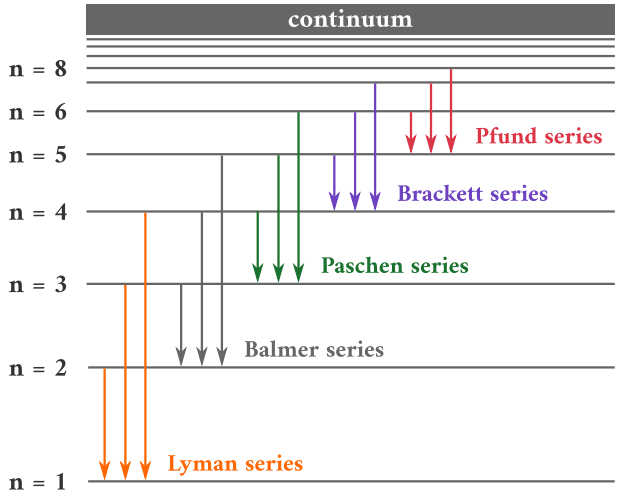


Figure 1.8.4 : Energy levels predicted by the Bohr model of hydrogen ($Z = 1$). (CC BY-NC; Ümit Kaya via LibreTexts)

These turn out to be the correct energy levels, apart from small corrections that cannot be accounted for in this pseudo-classical treatment. Despite the fact that the energies are essentially correct, the Bohr model masks the true quantum nature of the electron, which only emerges from a fully quantum mechanical analysis.

? Exercise 1.8.1

Calculate a value for the Bohr radius using Equation 1.8.8 to check that this equation is consistent with the value 52.9 pm. What would the radius be for $n = 1$ in the Li^{2+} ion?

? Exercise 1.8.2 : Rydberg states

How do the radii of the hydrogen orbits vary with n? Prepare a graph showing r as a function of n. States of hydrogen atoms with $n = 200$ have been prepared (called Rydberg states). What is the diameter of the atoms in these states?

The Wave Argument for Quantization

The above discussion is based off of a classical picture of an orbiting electron with the quantization from the angular momentum (Equation 1.8.7) requirement lifted from Planck's quantization arguments. Hence, only allows certain trajectories are stable (with

differing radii). However, as discussed previously, the electron will have a wavelike property also with a de Broglie wavelength λ

$$\lambda = \frac{h}{p}$$

Hence, a larger momentum p implies a shorter wavelength. That means as n increases (Equation 1.8.12), the wavelength must also increase; this is a common feature in quantum mechanics and will be often observed. In the Bohr atom, the circular symmetry and the wave property of the electron requires that the electron waves have an integer number of wavelengths (Figure 1.8.1*A*). If not, then the waves will overlap imperfectly and cancel out (i.e., the electron will cease to exist) as demonstrated in Figure 1.8.1*B*.

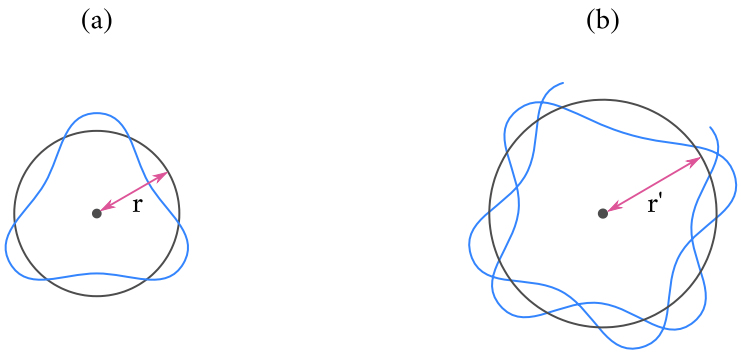


Figure 1.8.5 : Waves on a string have a wavelength related to the length of the string, allowing them to interfere constructively. (A) If we imagine the string bent into a closed circle, we get a rough idea of how electrons in circular orbits can interfere constructively. (B) If the wavelength does not fit into the circumference, the electron interferes destructively; it cannot exist in such an orbit. (CC BY-NC; Ümit Kaya via LibreTexts)

A more detailed discussion of the effect of electron waves in atoms will be discuss in the following chapters.

Derivation of the Rydberg Equation from Bohr Model

Given a prediction of the allowed energies of a system, how could we go about verifying them? The general experimental technique known as *spectroscopy* permits us to probe the various differences between the allowed energies. Thus, if the prediction of the actual energies, themselves, is correct, we should also be able to predict these differences. Let us assume that we are able to place the electron in Bohr's hydrogen atom into an energy state E_n for $n > 1$, i.e. one of its so-called *excited states*. The electron will rapidly return to its lowest energy state, known as the *ground state* and, in doing so, emit light. The energy carried away by the light is determined by the condition that the total energy is conserved (Figure 1.8.6).

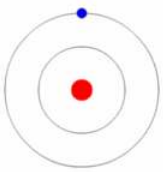


Figure 1.8.6 : A simple illustration of Bohr's model of the atom, with an electron making quantum leaps. (CC BY-SA 3.0 unported; Kurzon via Wikipedia)

Thus, if n_i is the integer that characterizes the initial (excited) state of the electron, and n_f is the final state (here we imagine that $n_f = 1$, but is applicable in cases that $n_f < n_i$, i.e., emission)

$$E_{nf} = E_{ni} - h\nu$$

or

$$\nu = \frac{E_{ni} - E_{nf}}{h} = \frac{Z^2 e^4 m_e}{8\epsilon_0^2 h^3}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \tag{1.8.13}$$

We can now identify the Rydberg constant R_H with the ratio of constants on the right hand side of Equation 1.8.13

$$R_H = \frac{m_e e^4}{8\epsilon_0^2 h^3}$$

Evaluating R_H from the fundamental constants in this formula gives a value within 0.5% of that obtained experimentally from the hydrogen atom spectrum.

Thus, by observing the emitted light, we can determine the energy difference between the initial and final energy levels, which results in the emission spectra discussed in Sections 1.4 and 1.5. Different values of n_f determine which emission spectrum is observed, and the examples shown in the figure are named after the individuals who first observed them. The figure below shows some of the transitions possible for different n_f and n_i values discussed previously.

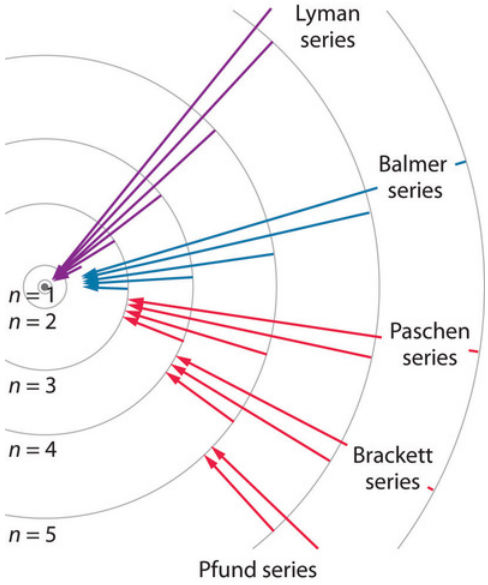


Figure 1.8.7 : Electron Transitions Responsible for the Various Series of Lines Observed in the Emission Spectrum of Hydrogen. The Lyman series of lines is due to transitions from higher-energy orbits to the lowest-energy orbit (n = 1); these transitions release a great deal of energy, corresponding to radiation in the ultraviolet portion of the electromagnetic spectrum. The Paschen, Brackett, and Pfund series of lines are due to transitions from higher-energy orbits to orbits with n = 3, 4, and 5, respectively; these transitions release substantially less energy, corresponding to infrared radiation. Orbits are not drawn to scale. (CC BY-NC-SA 3.0; anonymous)

If the atom absorbs light it ends up in an excited state as a result of the absorption. The absorption is only possible for light of certain frequencies, and again, conservation of energy determines what these frequencies are. If light is absorbed, then the final energy E_{nf} will be related to the initial energy E_{ni} with $n_f > n_i$ by

$$E_{nf} = E_{ni} + h\nu$$

or

$$\nu = \frac{E_{nf} - E_{ni}}{h} = \frac{Z^2 e^4 m_e}{8\epsilon_0^2 h^3}\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right)$$

Exercise 1.8.3

a. Calculate the energy of a photon that is produced when an electron in a hydrogen atom goes from an orbit with $n = 4$ to an orbit with $n = 1$.

b. What happens to the energy of the photon as the initial value of n approaches infinity?

Bohr's proposal explained the hydrogen atom spectrum, the origin of the Rydberg formula, and the value of the Rydberg constant. Specifically it demonstrated that the integers in the Rydberg formula are a manifestation of quantization. The energy, the angular momentum, and the radius of the orbiting electron all are quantized. This quantization also parallels the concept of stable orbits in the Bohr model. Only certain values of E, M, and r are possible, and therefore the electron cannot collapse onto the nucleus by continuously radiating energy because it can only have certain energies, and it cannot be in certain regions of space. The electron

can only jump from one orbit (quantum state) to another. The quantization means that the orbits are stable, and the electron cannot spiral into the nucleus in spite of the attractive Coulomb force.

Although Bohr's ideas successfully explained the hydrogen spectrum, they failed when applied to the spectra of other atoms. In addition a profound question remained. Why is angular momentum quantized in units of $\hbar$? As we shall see, de Broglie had an answer to this question, and this answer led Schrödinger to a general postulate that produces the quantization of angular momentum as a consequence. This quantization is not quite as simple as proposed by Bohr, and we will see that it is not possible to determine the distance of the electron from the nucleus as precisely as Bohr thought. In fact, since the position of the electron in the hydrogen atom is not at all as well defined as a classical orbit (such as the moon orbiting the earth) it is called an orbital. An electron orbital represents or describes the position of the electron around the nucleus in terms of a mathematical function called a wavefunction that yields the **probability** of positions of the electron.

Contributors and Attributions

- Mark Tuckerman (New York University)
- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")
- Template:Contriboundless

1.8: The Bohr Theory of the Hydrogen Atom is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.8: The Bohr Theory of the Hydrogen Atom** is licensed CC BY-NC-SA 4.0.

1.9: The Heisenberg Uncertainty Principle

Learning Objectives

- To understand that sometime you cannot know everything about a quantum system as demonstrated by the Heisenberg uncertainly principle.

In classical physics, studying the behavior of a physical system is often a simple task due to the fact that several physical qualities can be measured simultaneously. However, this possibility is absent in the quantum world. In 1927 the German physicist Werner Heisenberg described such limitations as the Heisenberg Uncertainty Principle, or simply the Uncertainty Principle, stating that it is not possible to measure both the momentum and position of a particle simultaneously.

The Heisenberg Uncertainty Principle is a fundamental theory in quantum mechanics that defines why a scientist cannot measure multiple quantum variables simultaneously. Until the dawn of quantum mechanics, it was held as a fact that all variables of an object could be known to exact precision simultaneously for a given moment. Newtonian physics placed no limits on how better procedures and techniques could reduce measurement uncertainty so that it was conceivable that with proper care and accuracy all information could be defined. Heisenberg made the bold proposition that there is a lower limit to this precision making our knowledge of a particle inherently uncertain.

Probability

Matter and photons are waves, implying they are spread out over some distance. What is the position of a particle, such as an electron? Is it at the center of the wave? The answer lies in how you measure the position of an electron. Experiments show that you will find the electron at some definite location, unlike a wave. But if you set up exactly the same situation and measure it again, you will find the electron in a different location, often far outside any experimental uncertainty in your measurement. Repeated measurements will display a statistical distribution of locations that appears wavelike (Figure 1.9.1).

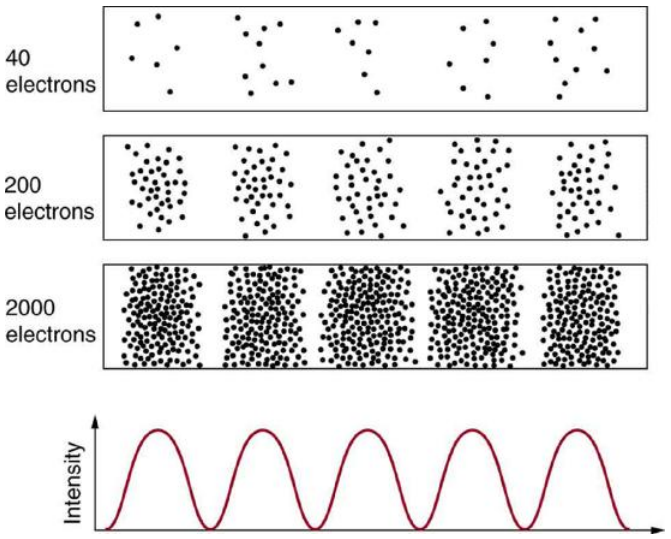


Figure 1.9.1 : The building up of the diffraction pattern of electrons scattered from a crystal surface. Each electron arrives at a definite location, which cannot be precisely predicted. The overall distribution shown at the bottom can be predicted as the diffraction of waves having the de Broglie wavelength of the electrons (CC BY 4.0; OpenStax).

After de Broglie proposed the wave nature of matter, many physicists, including Schrödinger and Heisenberg, explored the consequences. The idea quickly emerged that, *because of its wave character, a particle's trajectory and destination cannot be precisely predicted for each particle individually*. However, each particle goes to a definite place (Figure 1.9.1). After compiling enough data, you get a distribution related to the particle's wavelength and diffraction pattern. There is a certain *probability* of finding the particle at a given location, and the overall pattern is called a **probability distribution**. Those who developed quantum mechanics devised equations that predicted the probability distribution in various circumstances.

It is somewhat disquieting to think that you cannot predict exactly where an individual particle will go, or even follow it to its destination. Let us explore what happens if we try to follow a particle. Consider the double-slit patterns obtained for electrons and photons in Figure 1.9.2 . The interferrence patterns build up statistically as individual particles fall on the detector. This can be observed for photons or electrons—for now, let us concentrate on electrons. You might imagine that the electrons are interfering with one another as any waves do. To test this, you can lower the intensity until there is never more than one electron between the slits and the screen. The same interference pattern builds up!

This implies that a particle's probability distribution spans both slits, and the particles actually interfere with themselves. Does this also mean that the electron goes through both slits? An electron is a basic unit of matter that is not divisible. But it is a fair question, and so we should look to see if the electron traverses one slit or the other, or both. One possibility is to have coils around the slits that detect charges moving through them. What is observed is that an electron always goes through one slit or the other; it does not split to go through both.

But there is a catch. If you determine that the electron went through one of the slits, you no longer get a double slit pattern—instead, you get single slit interference. There is no escape by using another method of determining which slit the electron went through. Knowing the particle went through one slit forces a single-slit pattern. If you do not observe which slit the electron goes through, you obtain a double-slit pattern. How does knowing which slit the electron passed through change the pattern? The answer is fundamentally important—*measurement affects the system being observed*. Information can be lost, and in some cases it is impossible to measure two physical quantities simultaneously to exact precision. For example, you can measure the position of a moving electron by scattering light or other electrons from it. Those probes have momentum themselves, and by scattering from the electron, they change its momentum *in a manner that loses information*. There is a limit to absolute knowledge, even in principle.

Heisenberg's Uncertainty Principle

It is mathematically possible to express the uncertainty that, Heisenberg concluded, always exists if one attempts to measure the momentum and position of particles. First, we must define the variable "x" as the position of the particle, and define "p" as the momentum of the particle. The momentum of a photon of light is known to simply be its frequency, expressed by the ratio h/λ, where h represents Planck's constant and λ represents the wavelength of the photon. The position of a photon of light is simply its wavelength (λ). To represent finite change in quantities, the Greek uppercase letter delta, or Δ, is placed in front of the quantity. Therefore,

$$\Delta p = \frac{h}{\lambda} \tag{1.9.1}$$

$$\Delta x = \lambda \tag{1.9.2}$$

By substituting Δx for λ into Equation 1.9.1, we derive

$$\Delta p = \frac{h}{\Delta x} \tag{1.9.3}$$

or,

$$\Delta p \Delta x = h \tag{1.9.4}$$

early form of uncertainty principle

A Common Trend in Quantum Systems

Equation 1.9.4 can be derived by assuming the particle of interest is behaving as a particle, and not as a wave. Simply let $\Delta p = mv$, and $\Delta x = h/(mv)$ (from De Broglie's expression for the wavelength of a particle). Substituting in Δp for mv in the second equation leads to Equation 1.9.4.

Equation 1.9.4 was further refined by Heisenberg and his colleague Niels Bohr, and was eventually rewritten as

$$\Delta p_x \Delta x \geq \frac{h}{4\pi} = \frac{\hbar}{2} \tag{1.9.5}$$

with $\hbar = \frac{h}{2\pi} = 1.0545718 \times 10^{-34}\, m^2 \cdot kg/s$.

Equation 1.9.5 reveals that the more accurately a particle's position is known (the smaller Δx is), the less accurately the momentum of the particle in the x direction (Δp_x) is known. Mathematically, this occurs because the smaller Δx becomes, the larger Δp_x must become in order to satisfy the inequality. However, the more accurately momentum is known the less accurately position is known (Figure 1.9.2).

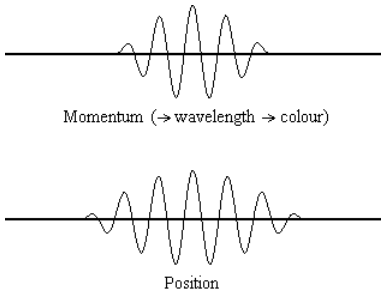


Figure 1.9.2 : The animation shows the relevant spreads in the uncertainty for position and momentum of light/photons (light wave's corresponding photon particle). From the result of de Broglie, we know that for a particle with known momentum, p will have a precise value for its de Broglie wavelength can be determined (and hence a specific color of the light).

What is the Proper Definition of Uncertainty?

Equation 1.9.5 relates the uncertainty of momentum and position. An immediate questions that arise is if Δx represents the full range of possible x values or if it is half (e.g., $\langle x \rangle \pm \Delta x$). Δx is the standard deviation and is a statistic measure of the spread of x values. The use of half the possible range is more accurate estimate of Δx. As we will demonstrated later, once we construct a wavefunction to describe the system, then both x and Δx can be explicitly derived. However for now, Equation 1.9.5 will work.

For example: If a problem argues a particle is trapped in a box of length, L, then the uncertainly of it position is $\pm L/2$. So the value of Δx used in Equation 1.9.5 should be $L/2$, not L.

Example 1.9.1

An electron is confined to the size of a magnesium atom with a 150 pm radius. What is the **minimum** uncertainty in its velocity?

Solution

The uncertainty principle (Equation 1.9.5):

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

can be written

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

and substituting $\Delta p = m \Delta v$ since the mass is not uncertain.

$$\Delta v \geq \frac{\hbar}{2\, m\, \Delta x}$$

the relevant parameters are

- mass of electron $m = m_e = 9.109383 \times 10^{-31}\ kg$
- uncertainty in position: $\Delta x = 150 \times 10^{-12} m$

$$\Delta v \geq \frac{1.0545718 \times 10^{-34}\ \cancel{kg}\, m^2/s}{(2)\ (9.109383 \times 10^{-31}\ \cancel{kg})\ (150 \times 10^{-12}\ \cancel{m})}$$

$$= 3.9 \times 10^5\ m/s$$

Exercise 1.9.1

What is the maximum uncertainty of velocity the electron described in Example 1.9.1 ?

Understanding the Uncertainty Principle through Wave Packets and the Slit Experiment

It is hard for most people to accept the uncertainty principle, because in classical physics the velocity and position of an object can be calculated with certainty and accuracy. However, in quantum mechanics, the wave-particle duality of electrons does not allow us to accurately calculate both the momentum and position because the wave is not in one exact location but is spread out over space. A "wave packet" can be used to demonstrate how either the momentum or position of a particle can be precisely calculated, but not both of them simultaneously. An accumulation of waves of varying wavelengths can be combined to create an average wavelength through an interference pattern: this average wavelength is called the "wave packet". The more waves that are combined in the "wave packet", the more precise the position of the particle becomes and the more uncertain the momentum becomes because more wavelengths of varying momenta are added. Conversely, if we want a more precise momentum, we would add less wavelengths to the "wave packet" and then the position would become more uncertain. Therefore, there is no way to find both the position and momentum of a particle simultaneously.



Several scientists have debated the Uncertainty Principle, including Einstein. Einstein created a slit experiment to try and disprove the Uncertainty Principle. He had light passing through a slit, which causes an uncertainty of momentum because the light behaves like a particle and a wave as it passes through the slit. Therefore, the momentum is unknown, but the initial position of the particle is known. Here is a video that demonstrates particles of light passing through a slit and as the slit becomes smaller, the final possible array of directions of the particles becomes wider. As the position of the particle becomes more precise when the slit is narrowed, the direction, or therefore the momentum, of the particle becomes less known as seen by a wider horizontal distribution of the light.

Example 1.9.2

The speed of a 1.0 g projectile is known to within $10^{-6}\ m/s$.

a. Calculate the *minimum* uncertainty in its position.
b. What is the *maximum* uncertainty of its position?

Solution a

From Equation 1.9.5, the $\Delta p_x = m\Delta v_x$ with $m = 1.0\ g$. Solving for Δx to get

$$\begin{aligned}\Delta x &= \frac{\hbar}{2m\Delta v} \\ &= \frac{1.0545718 \times 10^{-34}\ m^2 \cdot kg/s}{(2)(0.001\ kg)(10^{-6}\ m/s)} \\ &= 5.3 \times 10^{-26}\ m\end{aligned}$$

This negligible for all intents and purpose as expected for any macroscopic object.

Solution b

Unlimited (or the size of the universe). The Heisenberg uncertainty principles does not quantify the maximum uncertainty.

Exercise 1.9.2

Estimate the minimum uncertainty in the speed of an electron confined to a hydrogen atom within a diameter of $1 \times 10^{-10} m$?

Heisenberg's Uncertainty Principle not only helped shape the new school of thought known today as quantum mechanics, but it also helped discredit older theories. Most importantly, the Heisenberg Uncertainty Principle made it obvious that there was a fundamental error in the Bohr model of the atom. Since the position and momentum of a particle cannot be known simultaneously, Bohr's theory that the electron traveled in a circular path of a fixed radius orbiting the nucleus was obsolete. Furthermore, Heisenberg's uncertainty principle, when combined with other revolutionary theories in quantum mechanics, helped shape wave mechanics and the current scientific understanding of the atom.

Humor: Heisenberg and the Police

- Heisenberg get pulled over for speeding by the police. The officer asks him "Do you know how fast you were going?"
- Heisenberg replies, "No, but we know exactly where we are!"
- The officer looks at him confused and says "you were going 108 miles per hour!"
- Heisenberg throws his arms up and cries, "Great! Now we're lost!"

1.9: The Heisenberg Uncertainty Principle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.9: The Heisenberg Uncertainty Principle** by Kris Baumgartner, Sarah Woods is licensed CC BY 4.0.

1.E: The Dawn of the Quantum Theory (Exercises)

Solutions to select questions can be found online.

1.1A

Sodium metal has a threshold frequency of 4.40×10^{14} Hz. What is the kinetic energy of a photoelectron ejected from the surface of a piece of sodium when the ejecting photon is 6.20×10^{14} Hz? What is the velocity of this photoelectron? From which region of the electromagnetic spectrum is this photon?

1.1B

What is the longest-wavelength electromagnetic radiation that can eject a photoelectron from silver, given that the work function is 4.73 eV? Is this in the visible range?

1.1C

Find the longest-wavelength photon that can eject an electron from potassium, given that the work function is 2.24 eV. Is this visible electromagnetic radiation?

1.1C

What is the work function in eV of electrons in magnesium, if the longest-wavelength photon that can eject electrons is 337 nm?

1.1D

Calculate the work function in eV of electrons in aluminum, if the longest-wavelength photon that can eject the electromagnetic is 304 nm.

1.1E

What is the maximum kinetic energy in eV of electrons ejected from sodium metal by 450-nm electromagnetic radiation, given that the work function is 2.28 eV?

1.1F

UV radiation having a wavelength of 120 nm falls on gold metal, to which electrons are bound by 4.82 eV. What is the maximum kinetic energy of the ejected photoelectrons?

1.1G

Violet light of wavelength 400 nm ejects electrons with a maximum kinetic energy of 0.860 eV from sodium metal. What is the work function of electrons to sodium metal?

1.1H

UV radiation having a 300-nm wavelength falls on uranium metal, ejecting 0.500-eV electrons. What is the work function of electrons to uranium metal?

1.1I

What is the wavelength of electromagnetic radiation that ejects 2.00-eV electrons from calcium metal, given that the work function is 2.71 eV? What type of electromagnetic radiation is this?

1.1J

Find the wavelength of photons that eject 0.100-eV electrons from potassium, given that the work function is 2.24 eV. Are these photons visible?

1.1K

What is the maximum velocity of electrons ejected from a material by 80-nm photons, if they are bound to the material by 4.73 eV?

1.1L

Photoelectrons from a material with a work function of 2.71 eV are ejected by 420-nm photons. Once ejected, how long does it take these electrons to travel 2.50 cm to a detection device?

1.1M

A laser with a power output of 2.00 mW at a wavelength of 400 nm is projected onto calcium metal. (a) How many electrons per second are ejected? (b) What power is carried away by the electrons, given that the work function is 2.71 eV?

1.1N

(a) Calculate the number of photoelectrons per second ejected from a 1.00-mm 2 area of sodium metal by 500-nm electromagnetic radiation having an intensity of $1.30\ kW/m^2$ (the intensity of sunlight above the Earth's atmosphere). (b) Given that the work function is 2.28 eV, what power is carried away by the electrons? (c) The electrons carry away less power than brought in by the photons. Where does the other power go? How can it be recovered?

1.1O

Red light having a wavelength of 700 nm is projected onto magnesium metal to which electrons are bound by 3.68 eV. (a) Use $KE_e = h\nu - \Phi$ to calculate the kinetic energy of the ejected electrons. (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

1.1P

(a) What is the work function of electrons to a material from which 4.00-eV electrons are ejected by 400-nm electromagnetic radiation? (b) What is unreasonable about this result? (c) Which assumptions are unreasonable or inconsistent?

1.2A

a. Suppose the electron in a hydrogen atom is in the circular Bohr orbit with n = 30. How many times per second does it go around?
b. Suppose now the electron drops to the n = 29 state, emitting a single photon. What is the frequency of this photon, in cycles per second?
c. Comment on the relation between your answers in (a), (b) above. What would you guess the relation to be for n = 300?

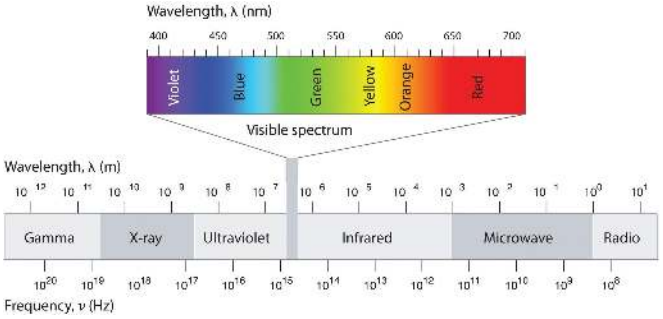
1.2B

The μ (muon) is a cousin of the electron, the only difference being its mass is 207 times greater. The μ has a lifetime of about 2 ms. If a beam of muons is directed at a solid, the muons will go into orbit around nuclei. The Bohr atom, with a muon replacing the electron, is a useful model for picturing this.

1. For a nucleus of charge Ze, how large is the n = 1 μ orbit compared with the electron orbit?
2. What is the frequency of the photon emitted by the μ in the n = 2 to n = 1 transition?
3. For the gold nucleus, the $n = 1$ μ orbit is inside the nucleus. Find the frequency of the emitted photon for $n = 2$ to $n = 1$ in this case. (Hint: you'll need the radius of the gold nucleus. Assume here that the positive charge is uniformly spread throughout the nucleus.)

1.3

Past Infrared region, in direction of the lower energies, the microwave region is located. In this region, radiation usually is characterized by frequency (ν) which is expressed in units of **MHz**, where **Hz** is a cycle per second. Given a microwave frequency of 2.0×10^4 **MHz**, calculate ν, λ, and energy per photon for this radiation and then compare the results with figure below.



1.4

Compare the Planck Distribution and the Rayleigh-Jean Distributions. For large values of ν, which one would be greater?

1.4

Plancks principal assumption was that energies of electronic oscillator can only have values $E = nh\nu$ and $\Delta E = h\nu$. In fact, as $v \to 0$ then $\Delta E \to 0$ and E becomes continuous. It should be expected that the nonclassical Plank distribution to go over to the classical Rayleigh-Jeans distribution at low frequencies, where $\Delta E \to 0$. Prove that Equation 1.2 reduces to Equation 1.1 as $v \to 0$.

Note: The Taylor expansion of an exponential

$$e^x \approx 1 + x + \left(\frac{x^2}{2!}\right) + ...$$

can truncated to $e^x \approx 1 + x$ when x is small.

1.5

The visible spectrum is in the 400-700 nm range, and contains about 40% of the sun's radiation intensity. Using the Planck Distribution, write an integral expression that can be evaluated to give this result (do not evaluate the integral).

1.13

What is the frequency and energy of a single 310 nm photon?

1.14

A laser emits 3.3×10^{17}photons per second. If the energy per photon is $6.4 \times 10^{-20} J$ per photon, find the a) wattage and b) wavelength of the laser. in what electromagnetic spectra is the laser in?

1.15

What is the max wavelength with a given temperature of 7500K?

Q1.15

A light bulb is a blackbody radiator. What temperature is required such that $\lambda_{max} = 400nm$?

1.16

An unknown elemental metal has work function of $\Phi = 8.01 \times 10^{-19} J$. Upon illumination with UV light of wavelength 162 nm, electrons are ejected with velocity of at $2.95 \times 10^3 \frac{m}{s}$. What is the threshold wavelength? What is the work function in units of eV? What metal does this correspond to (you will need to consult Table B1)?

1.16

Given the work function of sodium is 1.87 eV, find the kinetic energy of the ejected electrons when light of frequency 2.3 times greater than the threshold frequency is used to excite the electrons.

Q1.17

Find kinetic energy emitted off surface of tungsten that is radiated with radiation of 250 nm. Work function of tungsten 4.50 eV.

https://chem.libretexts.org/@go/page/198645

1.18

A smooth silver Thanksgiving platter and serving spoon (the pilgrims had knives and spoons but no forks) are irradiated with light of wavelength 317 nm. The work function is $\Phi = 6.825 \times 10^{-19} J$. What is the kinetic energy of the ejected electrons [eV]? The threshold frequency?

Q1.18

When a clean surface of silver is irradiated with light of wavelength 255 nm, the work function of ejected electrons is 4.18 eV. Calculate the kinetic energy in eV of the silver and the threshold frequency.

Q1.21

A line in the Paschen series of hydrogen has a wavelength of 1.01 x 10^{-6} m. Find the original energy of the electron.

1.22

How does the energy change when a particle absorbs and releases a photon? Show the effects on the state that the particle is in and the energy itself.

1.23

Show that the (a) wavelength of 100 nm occurs within the Lyman series, that (b) wavelength of 500 nm occurs within the Balmer series, and that (c) wavelength of 1000 nm occurs within the Paschen series. Identify the spectral regions to which these wavelengths correspond.

1.24

Calculate the wavelength and the energy of a photon associated with the series limit of the Balmer series.

1.25

For the following particles (a) an electron with a kinetic energy of 50eV, (b) a proton with a kinetic energy of 50eV , and (c) an electron in the second Bohr orbit of a hydrogen atom, calculate the de Broglie wavelength of each.

Q1.26

a. What is the velocity and wavelength of an electron with a voltage increase of 75 V?
b. What is the momentum of an electron with a de Broglie wavelength of 20 nm?(mass of an electron is $9.109 \times 10^{-31}\ kg$)

1.27

Through what potential a proton must initially at rest fall so its de Broglie wavelength is $1.83 \times 10^{-10}\ m$?

Q1.28

Calculate the energy and wavelength associated with a β particle that has fallen through a potential difference of 3.2 V. Take the mass of a β particle to be $9.1 \times 10^{-31}\mathrm{kg}$.

Q1.28

If a proton is going through a potential difference of 3.0V, what is the momentum and wavelength associated with this proton? (mass of a proton is equal to $1.6726 \times 10^{-27}\ kg$)

1.29

Neutron diffraction is a modern technique to study structure. In the neutron diffraction, a collimated beam of neutrons was generated at some temperature from high-energy neutron source. This is achieved at several accelerator facilities around the world. If the speed of neutron is $v_n = (3k_BT/m)^{1/2}$ with m as the mass of neutron. What is the required temperature so that neutrons have a de Broglie wavelength of 200 pm? The mass of a neutron to be $1.67 \times 10^{-27}\ kg$.

1.29

While studying quantum mechanics one day, you wondered what temperature would be required for the Jumbo Jawbreaker you were about to eat to have a de Broglie wavelength of 1.9×10^{-24} meters? Assuming that the speed of a Jumbo Jawbreaker can be calculated from the equation $\nu_n = (\frac{3k_BT}{m})\frac{1}{2}$. You quickly measure the mass of your Jumbo Jawbreaker and found it to be $0.1kg$.

1.30

For linear motion, show that a small change in the momentum, $\Delta \mathrm{p}$, changes a change in kinetic energy, $\Delta \mathrm{KE}$, of

$$\Delta \mathrm{KE} = \frac{p_0}{m}\Delta \mathrm{p}$$

where p_0 is initial momentum.

1.31

Derive the Bohr formula for $\frac{1}{\lambda_{vac}}$ for a multi-proton and single electron atom such as He^+ or Li^{2+}.

1.32

The series in the He^+ spectrum that corresponds to the set of transitions where the electron falls from a higher level into the n = 4 state is called the Pickering series, an important series in solar astronomy. Derive the formula for the wavelengths of the observed lines in this series. In what region of the spectrum does it occur?

1.33A

Using the Bohr model, find the third ionization energy for the Lithium atom in eV and in J.

1.33B

Find the ionization energy in eV and $kJ \cdot mol^{-1}$ of singly ionized helium in the $n = 3$ state using Bohr theory

1.34A

The speed of electron in an *n*th Bohr orbit is given by the equation:

$$v = \frac{e^2}{2\epsilon_0 nh}$$

The force acting between an electron and proton a distance r from one another is given by Coulomb's law:

$$f = \frac{e^2}{4\pi\epsilon_0 r^2}$$

The centrifugal force acts in opposition to the Coulombic force and is given by the equation:

$$f = \frac{mv^2}{r}$$

Find the values of v for the Bohr orbits of n = 4, n = 5, and n = 6, and find the total force in an atom between a proton and electron a distance of 5 x 10^{-11} m away from one another, with the electron moving at a speed of 2 x 10^6 m/s.

1.34B

Prove that the speed of electron in an *n*th Bohr orbit is $v = \frac{e^2}{2\epsilon_0 nh}$

Then find the first few values of v the Bohr orbit.

1.35

What is the uncertainty in an electron's position if the uncertainty in measuring its velocity is 5 $m \cdot s^{-1}$.

1.35

What is the *uncertainty* in the speed of an electron if we locate it to within 50 pm?

1.35

If we know the velocity of an electron to within $3.5 \times 10^7 \frac{m}{s}$, then what is the uncertainty in its position?

1.35

If a proton is located to within 1 angstrom, what is its uncertainty in velocity?

1.36

If the position of an electron is within the 10 pm interval, what is the uncertainty of the momentum? Is this value similar to that of an electron in the first Bohr orbit?

1.37

The Heisenberg Uncertainty Principle

$$\Delta x \Delta p \geq \frac{h}{4\pi}$$

Show both sides have the same units

1.38

The relationship between energy and time can be seen through the following uncertainty principle: $\Delta E \Delta t \geq h$. Through this relationship, it can be interpreted that a particle of mass m, the energy (E=mc^2) can come from nothing and return to nothing within a time $\Delta t \leq h/(mc^2)$. A real particle is one that lasts for time (Δt) or more; likewise, a particle that lasts for less than time (Δt) are called virtual particles. For a charged subatomic particle, a pion, the mass is $2.5 \times 10^{-28} kg$. For a pion to be considered a real particle, what is its minimum lifetime?

1.E: The Dawn of the Quantum Theory (Exercises) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

2: The Classical Wave Equation

The aim of this section is to give a fairly brief review of waves in various shaped elastic media—beginning with a taut string, then going on to an elastic sheet, a drumhead, first of rectangular shape then circular, and finally considering elastic waves on a spherical surface, like a balloon. The reason we look at this material here is that these are “real waves”, hopefully not too difficult to think about, and yet mathematically they are the solutions of the same wave equation the Schrödinger wave function obeys in various contexts, so should be helpful in visualizing solutions to that equation, in particular for the hydrogen atom.

2.1: The One-Dimensional Wave Equation
2.2: The Method of Separation of Variables
2.3: Oscillatory Solutions to Differential Equations
2.4: The General Solution is a Superposition of Normal Modes
2.5: A Vibrating Membrane
2.E: The Classical Wave Equation (Exercises)

2: The Classical Wave Equation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.1: The One-Dimensional Wave Equation

Learning Objectives

- To introduce the wave equation including time and position dependence

In the most general sense, waves are particles or other media with wavelike properties and structure (presence of crests and troughs).

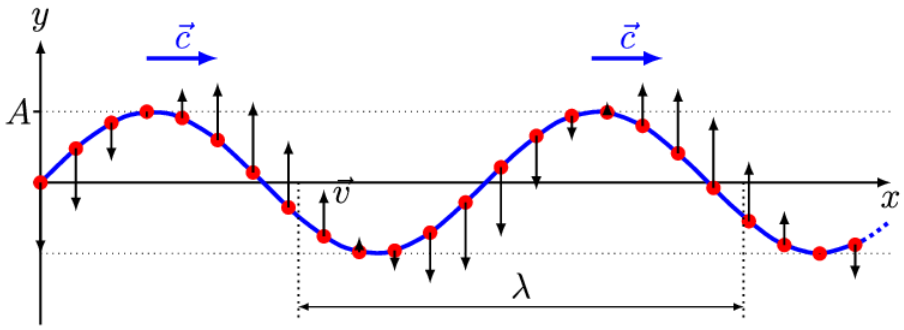


Figure 2.1.1 : A simple translational (transverse) wave. (CC BY-SA 4.0 International; And1mu via Wikimedia Commons)

The simplest wave is the (spatially) one-dimensional sine wave (Figure 2.1.1) with an varing amplitude A described by the equation:

$$A(x,t) = A_o \sin(kx - \omega t + \phi)$$

where

- A_o is the maximum amplitude of the wave, maximum distance from the highest point of the disturbance in the medium (the crest) to the equilibrium point during one wave cycle. In Figure 2.1.1 , this is the maximum vertical distance between the baseline and the wave.
- x is the space coordinate
- t is the time coordinate
- k is the wavenumber
- ω is the angular frequency
- ϕ is the phase constant.

One can categorize “waves” into two different groups: *traveling waves* and *stationary waves*.

Traveling Waves

Traveling waves, such as ocean waves or electromagnetic radiation, are waves that “move,” meaning that they have a frequency and are propagated through time and space. Another way of describing this property of “wave movement” is in terms of energy transmission – a wave travels, or transmits energy, over a set distance. The most important kinds of traveling waves in everyday life are electromagnetic waves, sound waves, and perhaps water waves, depending on where you live. It is difficult to analyze waves spreading out in three dimensions, reflecting off objects, etc., so we begin with the simplest interesting examples of waves, those restricted to move along a line. Let’s start with a rope, like a clothesline, stretched between two hooks. You take one end off the hook, holding the rope, and, keeping it stretched fairly tight, wave your hand up and back once. If you do it fast enough, you’ll see a single bump travel along the rope:

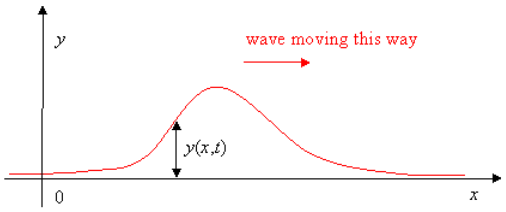


Figure 2.1.2 : A one-dimensional traveling wave at one instance of time t.

This is the simplest example of a *traveling wave*. You can make waves of different shapes by moving your hand up and down in different patterns, for example an upward bump followed by a dip, or two bumps. You'll find that the traveling wave *keeps the same shape* as it moves down the rope. Taking the rope to be stretched tightly enough that we can take it to be horizontal, we'll use its rest position as our *x*-axis (Figure 2.1.1). The y-axis is taken vertically upwards, and we only wave the rope in an up-and-down way, so actually $y(x,t)$ will be how far the rope is from its rest position at x at time t: that is, Figure 2.1.2 shows where the rope is at a single time t.

We can now express the observation that the wave "keeps the same shape" more precisely. Taking for convenience time $t = 0$ to be the moment when the peak of the wave passes $x = 0$, we graph here the rope's position at t = 0 and some later times t as a movie (Figure 2.1.3). Denoting the first function by $y(x,0) = f(x)$, then the second $y(x,t) = f(x - vt)$: it is the same function with the "same shape," but just moved over by vt, where v is the velocity of the wave.

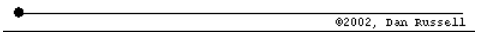


Figure 2.1.3 : A one-dimensional traveling wave at as a function of time. Traveling waves propagate energy from one spot to another with a fixed velocity v. (CC BY-NC-ND; Daniel A. Russell).

To summarize: on sending a traveling wave down a rope by jerking the end up and down, from observation the wave travels at constant speed and keeps its shape, so the displacement *y* of the rope at any horizontal position at x at time t has the form

$$y(x,t) = f(x - vt) \tag{2.1.1}$$

We are neglecting frictional effects—in a real rope, the bump gradually gets smaller as it moves along.

Standing Waves

In contrast to traveling waves, *standing waves*, or stationary waves, remain in a constant position with crests and troughs in fixed intervals. One way of producing a variety of standing waves is by plucking a melody on a set of guitar or violin strings. When placing one's finger on a part of the string and then plucking it with another, one has created a standing wave. The solutions to this problem involve the string oscillating in a sine-wave pattern (Figure 2.1.4) with no vibration at the ends. There is also no vibration at a series of equally-spaced points between the ends; these "quiet" places are *nodes*. The places of maximum oscillation are *antinodes*.

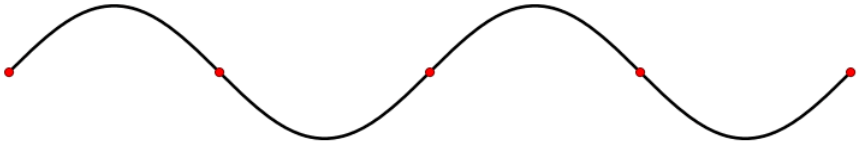


Figure 2.1.4 : Animation of standing wave in the stationary medium with marked wave nodes (red circles). (public domain; LucasVB).

Bound vs. Free particles and Traveling vs. Stationary Waves

Traveling waves exhibit movement and propagate through time and space and stationary wave have crests and troughs at fixed intervals separated by nodes. "Free" particles like the photoelectron discussed in the photoelectron effect, exhibit traveling

wave like properties. In contrast, electrons that are "bound" waves will exhibit stationary wave like properties. The latter was invoked for the Bohr atom for quantizing angular moment of an electron bound within a hydrogen atom.

The Wave Equation

The mathematical description of the one-dimensional waves (both traveling and standing) can be expressed as

$$\frac{\partial^2 u(x,t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 u(x,t)}{\partial t^2} \tag{2.1.2}$$

with u is the amplitude of the wave at position x and time t, and v is the velocity of the wave (Figure 2.1.2).

Equation 2.1.2 is called the classical wave equation in one dimension and is a *linear partial differential equation*. It tells us how the displacement u can change as a function of position and time and the function. The solutions to the wave equation ($u(x,t)$) are obtained by appropriate integration techniques. It may not be surprising that not all possible waves will satisfy Equation 2.1.2 and the waves that do must satisfy both **the initial conditions** and the **boundary conditions**, i.e. on how the wave is produced and what is happening on the ends of the string.

For example, for a standing wave of string with length L held taut at two ends (Figure 2.1.3), the boundary conditions are

$$u(0,t) = 0 \tag{2.1.3}$$

and

$$u(L,t) = 0 \tag{2.1.4}$$

for all values of t. As expected, different system will have different boundary conditions and hence different solutions.

Mathematical Origin of Quantization

The initial conditions and the boundary conditions used to solve the wave equation will result in restrictions of "allowed" waves to exist in a similar fashion that only certain solutions exist for the electrons in the Bohr atom.

The first six wave solutions $u(x,t)$ to Equation 2.1.2 subject to the boundary conditions in Equations 2.1.3 and 2.1.4 (discussed in detail later) results in the wave in Figure 2.1.5 . These are standing waves that exist with frequencies based on the number of *nodes* (0, 1, 2, 3,...) they exhibit (more discussed in the following Section).

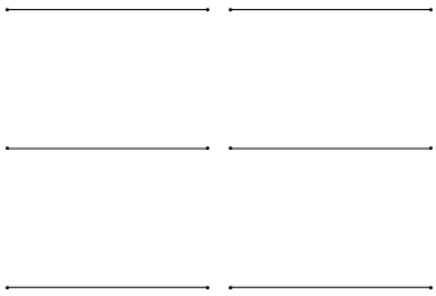


Figure 2.1.5 : Standing waves in a string (both spatially and temporally). The first six solutions ($u(x,t)$. from Wikipedia.

Curvature of Wave Solutions

Since the acceleration of the wave amplitude (right side of Equation 2.1.2) is proportional to $\frac{\partial^2}{\partial x^2}$, the greater curvature in the material produces a greater acceleration, i.e., greater changing velocity of the wave (Figure 2.1.4) and greater frequency of oscillation. As discussed later, the higher frequency waves (i..e, more nodes) are higher energy solutions; this as expected from the experiments discussed in Chapter 1 including Plank's equation $E = h\nu$.

Summary

Waves which exhibit movement and are propagated through time and space. The two basic types of waves are traveling and stationary. Both exhibit wavelike properties and structure (presence of crests and troughs) which can be mathematically described by a wavefunction or amplitude function. Both wave types display movement (up and down displacement), but in different ways.Traveling waves have crests and troughs which are constantly moving from one point to another as they travel over a length or distance. In this way, energy is transmitted along the length of a traveling wave. In contrast, standing waves have nodes at fixed positions; this means that the wave's crests and troughs are also located at fixed intervals. Therefore, standing waves only experience vibrational movement (up and down displacement) on these set intervals - no movement or energy travels along the length of a standing wave.

2.1: The One-Dimensional Wave Equation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.1: The One-Dimensional Wave Equation** by Michael Fowler is licensed CC BY-NC-SA 4.0.

2.2: The Method of Separation of Variables

Learning Objectives

- To be introduced to the *Separation of Variables* technique as method to solved wave equations

Solving the wave equation involves identifying the functions $u(x,t)$ that solve the partial differential equation that represent the amplitude of the wave at any position x at any time t

$$\frac{\partial^2 u(x,t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 u(x,t)}{\partial t^2} \tag{2.2.1}$$

This wave equation is a type of second-order partial differential equation (PDE) involving two variables - x and t. PDEs differ from ordinary differential equations (ODEs) that involve functions of only one variable. However, this difference makes PDEs appreciably more difficult to solve. In fact, the vast majority of PDE cannot be solved analytically and those classes of special PDEs that can be solved analytically invariably involve converting the PDE into one or more ODEs and then solving independently. One of these approaches is the the method of *separation of variables*.

Method of Separation of Variables

The general application of the *Method of Separation of Variables* for a wave equation involves three steps:

1. We find all solutions of the wave equation with the general form

   $$u(x,t) = X(x)T(t)$$

   for some function $X(x)$ that depends on x but not t and some function $T(t)$ that depends only on t, but not x. It is of course too much to expect that all solutions of Equation 2.2.1 are of this form, however, if we find a set of solutions $\{X_i(x)T_i(t)\}$ since the wave equation is a *linear equation*,

   $$u(x,t) = \sum_i c_i X_i(x) T_i(t) \tag{2.2.2}$$

   is also a solution for *any choice* of the constants c_i.
2. Impose constraints on the solutions based on the knowledge of the system. These are called the **boundary conditions**, which specify the values of $u(x,t)$ at the extremes ("boundaries"). This is a similar constraint to the solution as in initial value problems which the conditions $x(t_i)$ are specified at a specific time t_i. The goal is then to select the constants c_i in Equation 2.2.2 so that the boundary conditions are also satisfied.

Method of separation of variables is one of the most widely used techniques to solve partial differential equations and is based on the *assumption* that the solution of the equation is separable, that is, the final solution can be represented as a product of several functions, each of which is only dependent upon a single independent variable. If this assumption is incorrect, then clear violations of mathematical principles will be obvious from the analysis.

A Vibrating Spring Held Fixed Between Two Points

As discussed in Section 2.1, the solutions to the string example $u(x,t)$ for all x and t would be assumed to be a product of two functions: $X(x)$ and $T(t)$, where $X(x)$ is a function of only x, not t and $T(t)$ is a function of t, but not x.

$$u(x,t) = X(x)T(t) \tag{2.2.3}$$

Substitute Equation 2.2.3 into the one-dimensional wave equation (Equation 2.2.1) gives

$$\frac{\partial^2 X(x)T(t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 X(x)T(t)}{\partial t^2} \tag{2.2.4}$$

Since X is not a function of t and T is not a function of x, Equation 2.2.4 can be simplified

$$T(t)\frac{\partial^2 X(x)}{\partial x^2} = \frac{1}{v^2}X(x)\frac{\partial^2 T(t)}{\partial t^2} \tag{2.2.5}$$

Collecting the expressions that depend on x on the left side of Equation 2.2.5 and of t on the right side results in

$$\frac{1}{X(x)}\frac{\partial^2 X(x)}{\partial x^2} = \frac{1}{v^2}\frac{1}{T(t)}\frac{\partial^2 T(t)}{\partial t^2} \tag{2.2.6}$$

Equation 2.2.6 is an interesting equation since each side can be set to a fixed constant K as that is the only solution that works for all values of t and x. Therefore, the equation can be separated into two *ordinary differential equations:*

$$\frac{d^2 T(t)}{dt^2} - Kv^2 T(t) = 0 \tag{2.2.7}$$

$$\frac{d^2 X(x)}{dx^2} - KX(x) = 0 \tag{2.2.8}$$

Hence, by substituting the new product solution form (Equation 2.2.3) into the original wave equation (Equation 2.2.1), we converted a partial differential equation of two variables (x and t) into two *ordinary differential equations* (differential equation containing a function or functions of one independent variable and its derivatives). Each differential equation involves only one of the independent variables (x or t).

- If $K = 0$, then the solution is the trivial $u(x, y,) = 0$ solution (i.e., no wave exists).
- If $K > 0$, then the general solution of Equation 2.2.8 is

$$X(x) = Ae^{\sqrt{K}x} + Be^{-\sqrt{K}x} \tag{2.2.9}$$

At this stage, Equation 2.2.9 implies that the solution to the two ordinary differential wave equations will be an infinite number of waves with no quantization to limit those that are allowed (i.e., any values of A and B are possible). Narrowing down the general solution to a specific solution occurs when taking the boundary conditions into account.

The boundary conditions for this problem is that the wave amplitude equal to zero at the ends of the string

$$u(0, t) = X(x)T(t) = 0 \tag{2.2.10}$$

$$u(L, t) = X(x)T(t) = 0 \tag{2.2.11}$$

for all times t.

Applying the two boundary conditions in Equations 2.2.10 and 2.2.11 into the general solution in Equation 2.2.9 results into relationships between A and B:

$$X(x = 0) = A + B = 0 \tag{2.2.12}$$

and

$$X(x = L) = Ae^{\sqrt{K}L} + Be^{-\sqrt{K}L} = 0 \tag{2.2.13}$$

Ignore the Trivial Solution

One solution to this is that $A = B = 0$, but this is the trivial solution from $K = 0$ and one we ignore since it provides no physical solution to the problem other than the knowledge that $0 = 0$, which is not that inspiring of a result.

Both Equations 2.2.7 and 2.2.8 can be generalized into the following equations

$$\frac{d^2 y(x)}{dx^2} - k^2 y(x) = 0 \tag{2.2.14}$$

where k is a real constant (i.e., not complex). Equation 2.2.14 is a *homogeneous* second order linear differential equation. The general solution to these types of differential equations has the form

$$y(x) = e^{\alpha x} \tag{2.2.15}$$

where α is a constant to be determined by the constraints of system. Substituting Equation 2.2.15 into Equation 2.2.14 results in

$$\left(\alpha^2 - k^2\right) y(x) = 0 \tag{2.2.16}$$

For this equation to be satisfied, either

- $\alpha^2 - k^2 = 0$ or
- $y(x) = 0$.

The later is the trivial solution and is ignored and therefore

$$\alpha^2 - k^2 = 0 \tag{2.2.17}$$

so

$$\alpha = \pm k \tag{2.2.18}$$

Hence, there are two solutions to the general Equation 2.2.14, as expected for a second order differential equation (first order differential equations have one solution), which are a result from substituting the α values from Equation 2.2.18 into Equation 2.2.15

$$y(x) = e^{kx}$$

$$y(x) = e^{-kx}$$

The general solution can then be any linear combination of these two equations

$$y(x) = c_1 e^{kx} + c_2 e^{-kx} \tag{2.2.19}$$

Example 2.2.1 : General Solution

Solve

$$y'' + 3y' - 4y = 0$$

Solution

The strategy is to search for a solution of the form

$$y = e^{\alpha t}$$

The reason for this is that long ago some geniuses figured this stuff out and it works. Now calculate derivatives

$$y' = \alpha e^{\alpha t}$$

$$y'' = \alpha^2 e^{\alpha t}$$

Substituting into the differential equation gives

$$\alpha^2 e^{\alpha t} + 3(\alpha e^{\alpha t}) - 4(e^{\alpha t}) = 0$$
$$(\alpha^2 + 3\alpha - 4)e^{\alpha t} = 0$$

Now divide by $e^{\alpha t}$ to get

$$\alpha^2 + 3\alpha - 4 = 0$$
$$(\alpha - 1)(\alpha + 4) = 0$$
$$\alpha = 1$$

and

$$\alpha = -4$$

We can conclude that two solutions are

$$y_1 = e^t$$

and

$$y_2 = e^{-4t}$$

Now let

$$L(y) = y'' + 3y' - 4y$$

It is easy to verify that if y_1 and y_2 are solutions to

$$L(y) = 0$$

then

$$y = c_1 y_1 + c_2 y_2$$

is also a solution. More specifically we can conclude that

$$y = c_1 e^t + c_2 e^{-4t}$$

Represents a two dimensional family (vector space) of solutions. Later we will prove that this is the most general description of the solution space.

Example 2.2.2 : Boundary Conditions

Solve

$$y'' - y' - 6y = 0$$

with $y(0) = 1$ and $y'(0) = 2$.

Solution

As before we seek solutions of the form

$$y = e^{rt}$$

Now calculate derivatives

$$y' = re^{rt}$$
$$y'' = r^2 e^{rt}$$

Substituting into the differential equation gives

$$r^2 e^{rt} + (re^{rt}) - 6(e^{rt}) = 0$$
$$(r^2 - r - 6)e^{rt} = 0$$

Now divide by e^{rt} to get

$$r^2 - r - 6 = 0$$
$$(r-3)(r+2) = 0$$

We can conclude that two solutions are

$$y_1 = e^{3t}$$

and

$$y_2 = e^{-2t}$$

We can conclude that

$$y = c_1 e^{3t} + C_2 e^{-2t}$$

Represents a two dimensional family (a "vector space") of solutions. Now use the initial conditions to find that

$$1 = c_1 + c_2$$

We have that

$$y' = 3C_1 e^{3t} - 2C_2 e^{-2t}$$

Plugging in the initial condition with y', gives

$$2 = 3c_1 - 2c_2$$

This is a system of two equations and two unknowns. We can use linear algebra to arrive at

$$c_1 = \frac{4}{5}$$

and

$$C_2 = \frac{1}{5}$$

The final solution is

$$y = \frac{4}{5} e^{3t} + \frac{1}{5} e^{-2t}$$

When $K > 0$, the general solutions of Equations 2.2.7 and 2.2.8 are oscillatory in time and space, respectively, as discussed in the following section.

2.2: The Method of Separation of Variables is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.2: The Method of Separation of Variables** by Michael Fowler is licensed CC BY-NC-SA 4.0.

2.3: Oscillatory Solutions to Differential Equations

Learning Objectives

- Explore the basis of the oscillatory solutions to the wave equation
- Understand the consequences of boundary conditions on the possible solutions
- Rationalize how satisfying boundary conditions forces quantization (i.e., only solutions with specific wavelengths exist)

The boundary conditions for the string held to zero at both ends argue that $u(x,t)$ collapses to zero at the extremes of the string (Figure 2.3.1).



Figure 2.3.1 : Standing waves in a string (both spatially and temporally). from Wikipedia.

Unfortunately, when $K > 0$, the general solution (Equation 2.2.7) results in a sum of exponential decays and growths that cannot achieve the boundary conditions (except for the trivial solution); hence $K < 0$. This means we must introduce complex numbers due to the $\sqrt{K}$ terms in Equation 2.2.5. So we can rewrite K:

$$K = -p^2 \tag{2.3.1}$$

and Equation 2.2.4b can be

$$\frac{d^2X(x)}{dx^2} + p^2X(x) = 0 \tag{2.3.2}$$

The general solution to differential equations of the form of Equation 2.3.2 is

$$X(x) = Ae^{ipx} + Be^{-ipx} \tag{2.3.3}$$

Example 2.3.1

Verify that Equation 2.3.3 is the general form for differential equations of the form of Equation 2.3.2, which when substituted with Equation 2.3.1 give

$$X(x) = Ae^{ipx} + Be^{-ipx}$$

Solution

Expand the complex exponentials into trigonometric functions via Euler formula ($e^{i\theta} = \cos\theta + i\sin\theta$)

$$X(x) = A\left[\cos(px) + i\sin(px)\right] + B\left[\cos(px) - i\sin(px)\right]$$

collecting like terms

$$X(x) = (A+B)\cos(px) + i(A-B)\sin(px) \tag{2.3.4}$$

Introduce new *complex* constants $c_1 = A + B$ and $c_2 = i(A - B)$ so that the general solution in Equation 2.3.4 can be expressed as oscillatory functions

$$X(x) = c_1 \cos(px) + c_2 \sin(px) \tag{2.3.5}$$

Now let's apply the boundary conditions from Equation 2.2.7 to determine the constants c_1 and c_2. Substituting the first boundary condition ($X(x=0)=0$) into the general solutions of Equation 2.3.5 results in

$$\begin{aligned} X(x=0) = c_1 \cos(0) + c_2 \sin(0) &= 0 \\ c_1 + 0 &= 0 \\ c_1 &= 0 \end{aligned} \tag{2.3.6}$$

and substituting the second boundary condition ($X(x=L)=0$) into the general solutions of Equation 2.3.5 results in

$$X(x=L) = c_1 \cos(pL) + c_2 \sin(pL) = 0 \tag{2.3.7}$$

we already know that $c_1 = 0$ from the first boundary condition so Equation 2.3.7 simplifies to

$$c_2 \sin(pL) = 0 \tag{2.3.8}$$

Given the properties of sines, Equation 2.3.7 simplifies to

$$pL = n\pi \tag{2.3.9}$$

with $n = 0$ is the *trivial solution* that we ignore so $n = 1, 2, 3....$

$$p = \frac{n\pi}{L} \tag{2.3.10}$$

Substituting Equations 2.3.10 and 2.3.6 into Equation 2.3.5 results in

$$X(x) = c_2 \sin\left(\frac{n\pi x}{L}\right)$$

which can simplify to

$$X(x) = c_2 \sin(\omega x)$$

with

$$\omega = \frac{n\pi}{L}$$

A similar argument applies to the other half of the *ansatz* ($T(t)$).

? Exercise 2.3.1

Given two traveling waves:

$$\psi_1 = \sin(c_1 x + c_2 t) \quad \text{and} \quad \psi_2 = \sin(c_1 x - c_2 t)$$

a. Find the wavelength and the wave velocity of ψ_1 and ψ_2
b. Find the following and identify nodes:

$$\psi_+ = \psi_1 + \psi_2 \quad \text{and} \quad \psi_- = \psi_1 - \psi_2$$

Solution a

ψ_1 is a sin function. At every integer $n\pi$ where $n = 0, \pm 1, \pm 2, ...$, a sin function will be zero. Thus, $\psi_1 = 0$ when $c_1 x + c_2 t = \pi n$. Solving for the x, while ignoring trivial solutions:

$$x = \frac{n\pi - c_2 t}{c_1}$$

The velocity of this wave is:

$$\frac{dx}{dt} = -\frac{c_2}{c_1}$$

Similarly for ψ_2. At every integer $n\pi$ where $n = 0, \pm 1, \pm 2, ...$, a sin function will be zero. Thus, $\psi_2 = 0$ when $c_1 x - c_2 t = \pi n$. Solving for x, for ψ_2:

$$x = \frac{n\pi + c_2 t}{c_1}$$

The velocity of this wave is:

$$\frac{dx}{dt} = \frac{c_2}{c_1}$$

The wavelength for each wave is twice the distance between two successive nodes. In other words,

$$\lambda = 2(x_n - x_{n-1}) = \frac{2\pi}{c_1}$$

Solution b

Find $\psi_+ = \psi_1 + \psi_2$ and $\psi_- = \psi_1 - \psi_2$.

$$\begin{aligned}\psi_+ &= \sin(c_1 x + c_2 t) + \sin(c_1 x - c_2 t) \\ &= \sin(c_1 x)\cos(c_2 t) + \cancel{\cos(c_1 x)\sin(c_1 t)} + \sin(c_1 x)\cos(c_2 t) - \cancel{\cos(c_1 x)\sin(c_1 t)} \\ &= 2\sin(c_1 x)\cos(c_2 t)\end{aligned}$$

This should have a node at every $x = n\pi/c_1$ and

$$\begin{aligned}\psi_- &= \sin(c_1 x + c_2 t) - \sin(c_1 x - c_2 t) \\ &= \cancel{\sin(c_1 x)\cos(c_2 t)} + \cos(c_1 x)\sin(c_1 t) - \cancel{\sin(c_1 x)\cos(c_2 t)} + \cos(c_1 x)\sin(c_1 t) \\ &= 2\cos(c_1 x)\sin(c_2 t)\end{aligned}$$

2.3: Oscillatory Solutions to Differential Equations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.3: Oscillatory Solutions to Differential Equations** by Michael Fowler is licensed CC BY-NC-SA 4.0.

2.4: The General Solution is a Superposition of Normal Modes

Learning Objectives

- Separate the wave equation into individual spatial and temporal problems and solve them.
- Demonstrate that the general solution can be a superposition of solutions (normal modes)

As discussed previously, the solutions to the string example $u(x,t)$ for all x and t would be assumed to be a product of two functions: $X(x)$ and $T(t)$, where $X(x)$ is a function of only x, not t and $T(t)$ is a function of t, but not x.

$$u(x,t) = X(x)T(t)$$

By substituting the new product solution form into the original wave equation, one can obtain two *ordinary differential equations* (differential equation containing a function or functions of one independent variable and its derivatives). Each differential equation would involve only one of the independent variables (x or t).

Spatial Dependence of the Solution: $X(x)$

The boundary conditions for the string held to zero at both ends argue that $u(x,t)$ collapses to zero at the extremes of the string. Unfortunately, when $K > 0$, the general solution to the wave equation results in a sum of exponential decays and growths that cannot achieve the boundary conditions (except for the trivial solution that $u(x,t) = 0$); hence $K < 0$. This means we must introduce complex numbers due to the $\sqrt{K}$ terms in Equation 2.4.5. So we can rewrite K:

$$K = -p^2 \tag{2.4.1}$$

and Equation 2.4.4b can be

$$\frac{d^2X(x)}{dx^2} + p^2X(x) = 0 \tag{2.4.2}$$

The general solution to differential equations of the form of Equation 2.4.2 is

$$X(x) = Ae^{\sqrt{K}x} + Be^{-\sqrt{K}x} \tag{2.4.3}$$

that when substituted with Equation 2.4.1 give

$$X(x) = Ae^{ipx} + Be^{-ipx}$$

The complex exponentials in Equation 2.4.1 can be expressed as trigonometric functions via Euler formula ($e^{i\theta} = \cos\theta + i\sin\theta$)

$$X(x) = A\left[\cos(px) + i\sin(px)\right] + B\left[\cos(px) - i\sin(px)\right]$$

collecting like terms

$$X(x) = (A+B)\cos(px) + i(A-B)\sin(px) \tag{2.4.4}$$

The general solution in Equation 2.4.4 can be expressed as oscillatory functions by introducing new *complex* constants $C = A + B$ and $D = i(A - B)$:

$$X(x) = C\cos(px) + D\sin(px) \tag{2.4.5}$$

Exercise 2.4.1

Verify that Equation 2.4.3 is the general form for differential equations of the form of Equation 2.4.2.

Now let's apply the boundary conditions from Equation 2.4.5 to determine the constants C and D. Substituting the first boundary condition ($X(x = 0) = 0$) into the general solutions of Equation 2.4.5 results in

$$X(x=0)=0$$
$$C\cos(0)+D\sin(0)=0$$
$$C+0=0$$
$$C=0 \tag{2.4.7}$$

and substituting the second boundary condition ($X(x=L)=0$) into the general solutions of Equation 2.4.5 results in

$$X(x=L)=C\cos(pL)+D\sin(pL)=0 \tag{2.4.8}$$

we already know that $C=0$ from the first boundary condition so Equation 2.4.8 simplifies to

$$D\sin(pL)=0$$

Given the properties of sines, Equations 2.4.8 simplifies to

$$pL=n\pi$$

with $n=0$ is the *trivial solution* that we ignore so $n=1,2,3....$

$$p=\frac{n\pi}{L} \tag{2.4.9}$$

Substituting Equations 2.4.9 and 2.4.7 into Equation 2.4.5 results in

$$X(x)=D\sin\left(\frac{n\pi x}{L}\right) \tag{2.4.10}$$

Equation 2.4.10 presents a set of solutions to the spatial part of the solution to the wave equation subject to the boundary conditions (Figure 2.4.1). This set of solutions is infinitely large with individual solutions that are distinguished from each other by the n parameter introduced to account for the boundary conditions. This number is an example of a "quantum number" that are ubiquitously in quantum mechanics and are uniquely defined for each system.

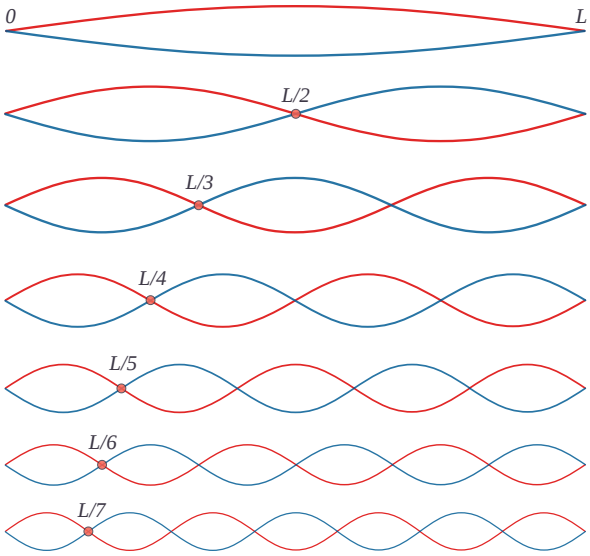


Figure 2.4.1: The first seven $X(x)$ solutions of a vibrating string of length L at the two times of maximum displacement (red and blue). (CC BY-NC; Ümit Kaya via LibreTexts).

Time Dependence of the Solution: $T(t)$

A similar argument applies to the other half of the *ansatz* ($T(t)$) originally proposed for the classical wave equation was obtain from solving Equation 2.4.4a, which qualitatively resembles the original spatial differential equation solved above (Equation 2.4.4b).

$$\frac{d^2T(t)}{dt^2}-Kv^2T(t)=0$$

However, the constraints extracted from solving spatial dependence apply to the time dependence. When Equations 2.4.1 and 2.4.9 are substituted into Equation 2.4.4a, a more simplified expression is obtained

$$\frac{d^2T(t)}{dt^2} + p^2v^2T(t) = \frac{d^2T(t)}{dt^2} + \left(\frac{nv\pi}{L}\right)^2 T(t) = 0 \tag{2.4.11}$$

Define a new constant: ω_n

$$\omega_n = \left(\frac{nv\pi}{L}\right)$$

and substitute into Equation 2.4.11

$$\frac{d^2T(t)}{dt^2} + \omega_n^2 T(t) = 0$$

This is the same functional form of Equation 2.4.2

$$T(t) = E\cos(\omega_n t) + F\sin(\omega_n t) \tag{2.4.12}$$

In contrast to the spatial dependence solution, we have no boundary conditions to use to identify the constants E and F.

The Principle of Superposition

Now let's revisit the original *ansatz* solution to the classical wave equation (Equation 2.4.1), which can be substituted with Equations 2.4.10 and 2.4.12

$$\begin{aligned} u(x,t) &= X(x)T(t) \\ &= \left[D\sin\left(\frac{n\pi x}{L}\right)\right](E\cos(\omega_n t) + F\sin(\omega_n t)) \end{aligned}$$

we can collect constants again with $G = DE$ and $H = DF$ and introduce a n dependence to each since E and F may be n dependence.

$$u_n(x,t) = \underbrace{[G_n\cos(\omega_n t) + H_n\sin(\omega_n t)]}_{\text{temporal solution}}\underbrace{\sin\left(\frac{n\pi x}{L}\right)}_{\text{spatial solution}} \tag{2.4.13}$$

The functions represented in Equation 2.4.13 are set of solutions including both spatial and temporal features that solve the wave equation of a string held tight on two ends.

Linearity of the Wave Equation

The wave equation has a very important property: if we have two solutions to the equation, then the sum of the two is *also* a solution to the equation. It's easy to check this:

$$\frac{\partial^2(f+g)}{\partial x^2} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 g}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2} + \frac{1}{v^2}\frac{\partial^2 g}{\partial t^2} = \frac{1}{v^2}\frac{\partial^2(f+g)}{\partial t^2}$$

Any differential equation for which this property holds is called *a linear differential equation*. Also note that $af(x,t) + bg(x,t)$ is also a solution to the equation if a, b are constants. So you can add together—superpose—multiples of any two solutions of the wave equation to find a new function satisfying the equation.

The linearity property is easy to interpret *visually*: if you can draw two wave solutions, then at each point on the string simply add the displacement $u_n(x,t)$ of one wave to the other $u_m(x,t)$—the sum of the two waves together is a solution. So, for example, as two traveling waves moving along the string in opposite directions meet each other, the displacement of the string at any point at any instant is just the sum of the displacements it would have had from the two waves singly.



Figure 2.4.2: Interference of Multiple Pulses. ("for non-profit educational purposes."; Tom Walsh via ophysics.com)

This simple addition of the displacements is interference, doubtless because if the waves meeting have displacements in opposite directions, the string will be displaced less than by a single wave. This is also called the ***Principle of Superposition***.

Principle of Superposition

The *Principle of Superposition* is the sum of two or more solutions is also a solution.

Since the wave equation is a linear homogeneous differential equation, the total solution can be expressed as a sum of all possible solutions described by Equation 2.4.13.

$$u(x,t) = \sum_{n=1}^{\infty} u_n(x,t) \tag{2.4.14}$$

$$= \sum_{n=1}^{\infty} (G_n \cos(\omega_n t) + H_n \sin(\omega_n t)) \sin\left(\frac{n\pi x}{L}\right) \tag{2.4.15}$$

Each $u_n(x,t)$ solution is called a **normal mode** of the system and can be characterized via their corresponding frequencies $\frac{n\pi}{L}$ with $n = 1, 2, 3...$. The spatial dependence of the first seven normal modes are shown in Figure 2.4.1 and are standing waves. The first term with $n = 1$ is typically called the *fundamental* and each subsequent modes is called an *overtone* or *harmonic*. The temporal dependence of the normal modes is sinusoidal with angular frequencies ω_n that can be expanded to natural frequencies ν_n via

$$\nu_n = \frac{\omega_n}{2\pi} = \frac{nv}{2L} \tag{2.4.16}$$

Hence, as the spatial curvature of the normal mode increases, the temporal oscillation of that mode also increases. This is a common trait in quantum mechanical systems and is a direct consequence of the wave equation.

2.4: The General Solution is a Superposition of Normal Modes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.4: The General Solution is a Superposition of Normal Modes** by Michael Fowler is licensed CC BY-NC-SA 4.0.

2.5: A Vibrating Membrane

Learning Objectives

- To apply the wave equations to a two-dimensional membrane (rectangles and circles)
- To recognize the possible geometries of a nodes in two-dimensional systems

So far, we've looked at waves in one dimension, traveling along a string or sound waves going down a narrow tube. However, waves in higher dimensions than one are very familiar—water waves on the surface of a pond, or sound waves moving out from a source in three dimensions. It is pleasant to find that these waves in higher dimensions satisfy wave equations which are a very natural extension of the one we found for a string, and—very important—they also satisfy the *Principle of Superposition*, in other words, if waves meet, you just add the contribution from each wave. In the next two paragraphs, we go into more detail, but this Principle of Superposition is the crucial lesson.

The Wave Equation and Superposition in More Dimensions

What happens in higher dimensions? Let's consider two dimensions, for example waves in an elastic sheet like a drumhead. If the rest position for the elastic sheet is the $(x,\ y)$ plane, so when it's vibrating it's moving up and down in the z-direction, its configuration at any instant of time is a function. $u(x,y,t)$

In fact, we could do the same thing we did for the string, looking at the total forces on a little bit and applying Newton's Second Law. In this case that would mean taking one little bit of the drumhead, and instead of a small stretch of string with tension pulling the two ends, we would have a small *square* of the elastic sheet, with tension pulling all around the edge. Remember that the net force on the bit of string came about because the string was curving around, so the tensions at the opposite ends tugged in slightly different directions, and did not cancel. The $\frac{\partial^2}{\partial x^2}$ term measured that curvature, the rate of change of slope. In two dimensions, thinking of a small square of the elastic sheet, things are more complicated. Visualize the bit of sheet to be momentarily like a tiny patch on a balloon, you'll see it curves in two directions, and tension forces must be tugging all around the edges. The total force on the little square comes about because the tension forces on opposite sides are out of line if the surface is curving around, now we have to add *two* sets of almost-opposite forces from the two pairs of sides. The math is now shown here, but it is at least plausible that the equation is:

$$\frac{\partial^2 u(x,y,t)}{\partial x^2}+\frac{\partial^2 u(x,y,t)}{\partial y^2}=\frac{1}{v^2}\frac{\partial^2 u(x,y,t)}{\partial t^2} \tag{2.5.1}$$

The physics of this equation is that the acceleration of a tiny bit of the sheet comes from out-of-balance tensions caused by the sheet curving around in *both* the *x*- and *y*-directions, this is why there are the two terms on the left hand side.

And, going to three dimensions is easy: add one more term to give

$$\frac{\partial^2 u(x,y,,z,t)}{\partial x^2}+\frac{\partial^2 u(x,y,z,t)}{\partial y^2}+\frac{\partial^2 u(x,y,z,t)}{\partial z^2}=\frac{1}{v^2}\frac{\partial^2 u(x,y,z,t)}{\partial t^2} \tag{2.5.2}$$

This sum of partial differentiations in space is so common in physics that there's a shorthand:

$$\nabla^2=\left(\frac{\partial^2}{\partial x^2},\frac{\partial^2}{\partial y^2},\frac{\partial^2}{\partial z^2}\right)$$

so Equation 2.5.2 can be more easily written as

$$\nabla^2 u(x,y,z,t)=\frac{1}{v^2}\frac{\partial^2 u(x,y,z,t)}{\partial t^2}$$

Just as we found in one dimension traveling harmonic waves (no boundary conditions)

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

with $\omega=\nu k$, you can verify that the *three*-dimensional equation has harmonic solutions

$$u(x,y,z,t)=A\sin(k_x x+k_x+k_z z-\omega t)$$

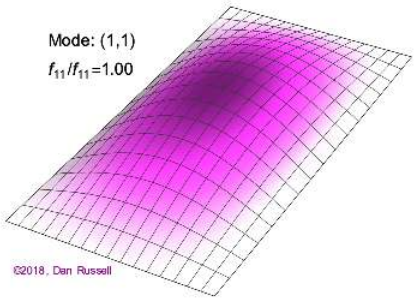
with $\omega = \nu|\vec{k}|$ where $|k| = \sqrt{k_x^2 + k_y^2 + k_z^2}$.

$\vec{k}$ is a vector in the direction the wave is moving. The electric and magnetic fields in a radio wave or light wave have just this form (or, closer to the source, a very similar equivalent expression for outgoing spheres of waves, rather than plane waves).

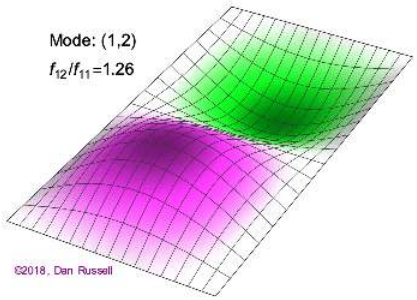
> It's important to realize that the 2D wave equation (Equation 2.5.1) is still a ***linear*** equation, so the Principle of Superposition still holds. If two waves on an elastic sheet, or the surface of a pond, meet each other, the result at any point is given by simply adding the displacements from the individual waves. We'll begin by thinking about waves propagating freely in two and three dimensions, than later consider waves in restricted areas, such as a drum head.

Vibrational Modes of a Rectangular Membrane

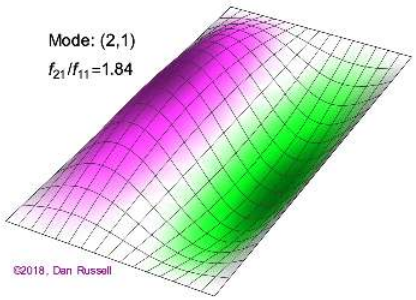
A one-dimensional wave does not have a choice in how it propagates: it just moves along the line (well, it could get partly reflected by some change in the line and part of it go backwards). However, when we go to higher dimensions, how a wave disturbance starting in some localized region spreads out is far from obvious. But we can begin by recalling some simple cases: dropping a pebble into still water causes an outward moving circle of ripples. If we grant that light is a wave, we notice a beam of light changes direction on going from air into glass. Of course, it is not immediately evident that light *is* a wave: we'll talk a lot more about that later. A few solutions (both temporal and spatials) are shown below together with their quantum numbers (n_x and n_y).



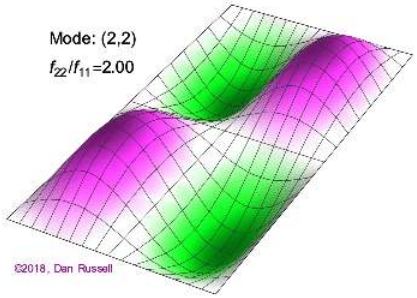
The ($n_x = 1$, $n_y = 1$) Solution



The ($n_x = 1$, $n_y = 2$) Solution



The ($n_x = 2$, $n_y = 1$) Solution



The ($n_x = 2$, $n_y = 2$) Solution

Figure 2.5.1: Select Vibrational Modes of a Rectangular Membrane. Animations courtesy of Dr. Dan Russell via Grad. Prog. Acoustics, Penn State).

Solving for the function $u(x, y, t)$ in a vibrating, rectangular membrane is done in a similar fashion by separation of variables, and setting boundary conditions. The solved function is very similar, where

$$u(x,y,t) = A_{nm}\cos(\omega_{nm}t + \phi_{nm})\sin\left(\frac{n_x\pi x}{a}\right)\sin\left(\frac{n_y\pi y}{b}\right) \tag{2.5.3}$$

where

- a is the length of the rectangular membrane and b is the width, and
- n_x and n_y are two quantum numbers (one in each dimension).

As with the solutions to 1-D wave equations, a **node** is a point (or line) on a structure that does not move while the rest of the structure is vibrating.

Example 2.5.1: Nodal Geometries in Rectangular Membranes

For the following 2-D solutions in Figure **2.5.1**, how many nodes are there, what is the geometry and how would you characterize them?

a. $(n_x = 1, n_y = 1$
b. $(n_x = 2, n_y = 1$

Solution a

The $(n_x = 1, n_y = 1$ solution in Figure **2.5.1** has zero nodes. That is, no spot in the membrane (other than the boundaries) is not moving during the motion of the membrane.

Solution b

The $(n_x = 2, n_y = 1$ solution in Figure **2.5.1** has one node. It is a line at half the length of the x direction and extends over the entire length of the y direction.

Exercise 2.5.1

For the $(n_x = 2, n_y = 2$ solution to a rectangular membrane in Figure 2.5.1 : how many nodes are there, what is the geometry and how would you characterize them?

Vibrational Modes of a Circular Membrane

The basic principles of a vibrating rectangular membrane applies to other 2-D members including circular membranes. However, the mathematics and solutions are a bit more complicated. The solutions are best represented in polar notation (instead of rectangular like in Equation 2.5.3) and have the following functional form

$$u(r,\theta,t) = J_m(\lambda_{mn}r)\cos m\theta\cos c\lambda_{mn}t$$

where J_m are Bessel functions (these are oscillatory functions) and λ are constants. This system has two quantum numbers (m and n) that serve the same function as n_x and n_y do in the rectangular membranes. In the animations in Figure 2.5.2 , the nodal diameters and circles show up as white regions that do not oscillate, while the red and blue regions indicate positive and negative displacements.

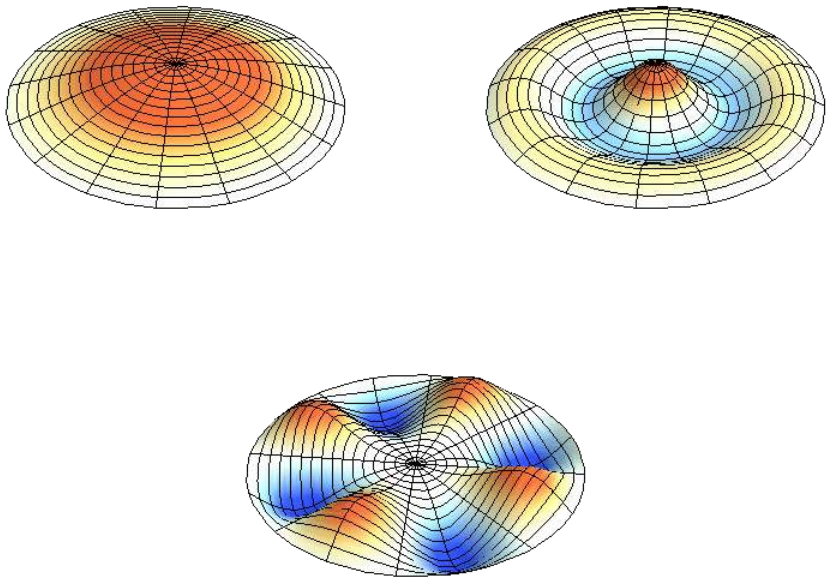


Figure **2.5.2**: Select Vibrational Modes of a Circular Membrane. (CC BY-SA-ND 4.0; Dan Russell, via Grad. Prog. Acoustics, Penn State.

Figure **2.5.2** (left) shows the fundamental mode shape for a vibrating circular membrane, while the other two modes are excited modes with more complex nodal character.

? Exercise 2.5.2

How many nodes are there in the three solutions for a circular membrane in Figure **2.5.2**, what are their geometries, and how would you characterize them?

2.5: A Vibrating Membrane is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.5: A Vibrating Membrane** by Delmar Larsen, Michael Fowler is licensed CC BY-NC-SA 4.0.

2.E: The Classical Wave Equation (Exercises)

Solutions to select questions can be found online.

2.1A

Find the general solutions to the following differential equations:

a. $\frac{d^2y}{dx^2} - 4y = 0$
b. $\frac{d^2y}{dx^2} - 3\frac{dy}{dx} - 54y = 0$
c. $\frac{d^2y}{dx^2} + 9y = 0$

2.1B

Find the general solutions to the following differential equations:

a. $\frac{d^2y}{dx^2} - 16y = 0$
b. $\frac{d^2y}{dx^2} - 6\frac{dy}{dx} + 27y = 0$
c. $\frac{d^2y}{dx^2} + 100y = 0$

2.1C

Find the general solutions to the following differential equations:

a. $\frac{dy}{dx} - 4\sin(x)y = 0$
b. $\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + 6y = 0$
c. $\frac{d^2y}{dx^2} = 0$

2.2A

Practice solving these first and second order homogeneous differential equations with given boundary conditions:

a. $\frac{dy}{dx} = ay$ with $y(0) = 11$
b. $\frac{d^2y}{dt^2} = ay$ with $y(0) = 6$ and $y'(0) = 4$
c. $\frac{d^2y}{dt^2} + \frac{dy}{dt} - 42y = 0$ with $y(0) = 2$ and $y'(0) = 0$

2.3A

Prove that $x(t)$ = $\cos(\theta)$ oscillates with a frequency

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$$

Prove that $x(t)$ = $\cos(\theta)$ also has a period

$$T = 2\pi\sqrt{\frac{m}{k}}$$

where k is the force constant and m is mass of the body.

2.3B

Try to show that

$$x(t) = \sin(\omega t)$$

oscillates with a frequency

$$\nu = \omega/2\pi$$

Explain your reasoning. Can you give another function of x(t) that have the same frequency.

2.3C

Which two functions oscillate with the same frequency?

a. $x(t) = \cos(\omega t)$
b. $x(t) = \sin(2\omega t)$
c. $x(t) = A\cos(\omega t) + B\sin(\omega t)$

2.3D

Prove that $x(t) = \cos(\omega(t))$ oscillates with a frequency

$$\nu = \frac{\omega}{2\pi}.$$

Prove that $x(t) = A\cos(\omega(t)) + B\sin(\omega(t))$ oscillates with the same frequency:

$$\nu = \frac{\omega}{2\pi}.$$

2.4

Show that the differential equation:

$$\frac{d^2y}{dx^2} + y(x) = 0$$

has a solution

$$y(x) = 2\sin x + \cos x$$

2.7

For a classical harmonic oscillator, the displacement is given by

$$\xi(t) = v_0\sqrt{\frac{m}{k}}\sin\sqrt{\frac{k}{m}}t$$

where $\xi = x - x_0$. Derive an expression for the velocity as a function of time, and determine the times at which the velocity of the oscillator is zero.

2.11

Verify that

$$Y(x,t) = A\sin\left(\frac{2\pi}{\lambda}(x - vt)\right)$$

has a frequency $\nu = v/\lambda$ and wavelength λ traveling right with a velocity v.

2.13A

Explain (in words) how to expand the Hamiltonian into two dimensions and use it solve for the energy

2.13B

Given that the Schrödinger equation for a two-dimensional box, with sides a and b, is

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{(8\pi^2 mE)}{h^2}\Psi(x,y) = 0$$

and it has the boundary conditions of

$$\Psi(0,y) = \Psi(a,y) = 0 \text{ and } \Psi(o,x) = \Psi(x,b) = 0$$

for all x and y values, show that

$$E_{2,2} = \left(\frac{h^2}{2ma^2}\right) + \left(\frac{h^2}{2mb^2}\right).$$

2.14

Explain, in words, how to expand the Schrödinger Equations into a three-dimensional box

2.18

Solving for the differential equation for a pendulum gives us the following equation,

$$\phi(x) = c_1 \cos\sqrt{\frac{g}{L}} + c_2 \sin\sqrt{\frac{g}{L}}$$

Assuming $c_1 = 2$, $c_3 = 5$, $g = 7$ and $L = 3$, what is the position of the pendulum initially? Does this make sense in the real world. Why or why not? (We can ignore units for this problem).

2.23

Consider a Particle of mass m in a one-dimensional box of length a. Its average energy is given by

$$\langle E \rangle = \frac{1}{2m}\langle p^2 \rangle$$

Because

$$\langle p \rangle = 0$$

$$\langle p^2 \rangle = \sigma_p^2$$

where σ_p can be called the uncertainty in p. Using the Uncertainty Principle, show that the energy must be at least as large as $\hbar/8ma^2$ because σ_x, the uncertainty in x, cannot be larger than a.

2.33

Prove $y(x,t) = A\cos[2\pi/\lambda(x - vt)]$ is a wave traveling to the right with velocity v, wavelength λ, and period λ/v.

2.E: The Classical Wave Equation (Exercises) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

3: The Schrödinger Equation & a Particle in a Box

The particle in a box model (also known as the infinite potential well or the infinite square well) describes a particle free to move in a small space surrounded by impenetrable barriers. The model is mainly used as a hypothetical example to illustrate the differences between classical and quantum systems. In classical systems, for example a ball trapped inside a large box, the particle can move at any speed within the box and it is no more likely to be found at one position than another. However, when the well becomes very narrow (on the scale of a few nanometers), quantum effects become important. The particle may only occupy certain positive energy levels. The particle in a box model provides one of the very few problems in quantum mechanics which can be solved analytically, without approximations. This means that the observable properties of the particle (such as its energy and position) are related to the mass of the particle and the width of the well by simple mathematical expressions. Due to its simplicity, the model allows insight into quantum effects without the need for complicated mathematics. It is one of the first quantum mechanics problems taught in undergraduate physics courses, and it is commonly used as an approximation for more complicated quantum systems.

3.1: The Schrödinger Equation
3.2: Linear Operators in Quantum Mechanics
3.3: The Schrödinger Equation is an Eigenvalue Problem
3.4: The Quantum Mechanical Free Particle
3.5: Wavefunctions Have a Probabilistic Interpretation
3.6: The Energy of a Particle in a Box is Quantized
3.7: Wavefunctions Must Be Normalized
3.8: The Average Momentum of a Particle in a Box is Zero
3.9B: Particle in a Finite Box and Tunneling (optional)
3.9: The Uncertainty Principle Redux - Estimating Uncertainties from Wavefunctions
3.10: A Particle in a Two-Dimensional Box
3.11: A Particle in a Three-Dimensional Box
3.E: The Schrödinger Equation and a Particle in a Box (Exercises)

Thumbnail: The quantum wavefunction of a particle in a 2D infinite potential well of dimensions L_x *and* L_y*. The wavenumbers are* $n_x = 2$ *and* $n_y = 2$*. Image used with permission (Public Domain; Inductiveload).*

3: The Schrödinger Equation & a Particle in a Box is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.1: The Schrödinger Equation

Learning Objectives

- To be introduced to the general properties of the Schrödinger equation and its solutions.

De Broglie's doctoral thesis, defended at the end of 1924, created a lot of excitement in European physics circles. Shortly after it was published in the fall of 1925 Pieter Debye, Professor of Theoretical Physics at Zurich and Einstein's successor, suggested to Erwin Schrödinger that he give a seminar on de Broglie's work. Schrödinger gave a polished presentation, but at the end Debye remarked that he considered the whole theory rather childish: why should a wave confine itself to a circle in space? It wasn't as if the circle was a waving circular string, real waves in space diffracted and diffused, in fact they obeyed three-dimensional wave equations, and that was what was needed. This was a direct challenge to Schrödinger, who spent some weeks in the Swiss mountains working on the problem and constructing his equation. There is no rigorous derivation of Schrödinger's equation from previously established theory, but it can be made very plausible by thinking about the connection between light waves and photons, and construction an analogous structure for de Broglie's waves and electrons (and later, other particles).

The Schrödinger Equation: A Better Approach

While the Bohr model is able to predict the allowed energies of any single-electron atom or cation, it by no means, a general approach. Moreover, it relies heavily on classical ideas, clumsily grafting quantization onto an essentially classical picture, and therefore, provides no real insights into the true quantum nature of the atom. Any rule that might be capable of predicting the allowed energies of a quantum system must also account for the *wave-particle duality* and implicitly include a wave-like description for particles. Nonetheless, we will attempt a heuristic argument to make the result at least plausible. In classical electromagnetic theory, it follows from Maxwell's equations that each component of the electric and magnetic fields in vacuum is a solution of the 3-D wave equation for electromagnetic waves:

$$\nabla^2\Psi(x,y,z,t) - \frac{1}{c^2}\frac{\partial^2\Psi(x,y,z,t)}{\partial t^2} = 0 \tag{3.1.1}$$

The wave equation in Equation 3.1.1 is the three-dimensional analog to the wave equation presented earlier (Equation 2.1.1) with the velocity fixed to the known speed of light: c. Instead of a partial derivative $\frac{\partial^2}{\partial x^2}$ in one dimension, the Laplacian (or "del-squared") operator is introduced:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{3.1.2}$$

Corresponding, the solution to this 3D equation wave equation is a function of *four* independent variables: x, y, z, and t and is generally called the **wavefunction** ψ.

We will attempt now to create an analogous equation for de Broglie's matter waves. Accordingly, let us consider a only 1-dimensional wave motion propagating in the *x*-direction. At a given instant of time, the form of a wave might be represented by a function such as

$$\Psi(x) = f\left(\frac{2\pi x}{\lambda}\right) \tag{3.1.3}$$

where $f(\theta)$ represents a sinusoidal function such as $\sin\theta$, $\cos\theta$, $e^{i\theta}$, $e^{-i\theta}$ or some linear combination of these. The most suggestive form will turn out to be the complex exponential, which is related to the sine and cosine by Euler's formula

$$e^{\pm i\theta} = \cos\theta \pm i\sin\theta \tag{3.1.4}$$

Each of the above is a periodic function, its value repeating every time its argument increases by 2π. This happens whenever x increases by one wavelength λ. At a fixed point in space, the time-dependence of the wave has an analogous structure:

$$T(t) = f(2\pi\nu t) \tag{3.1.5}$$

where ν gives the number of cycles of the wave per unit time. Taking into account both x and t dependence, we consider a wavefunction of the form

$$\Psi(x,t) = exp\left[2\pi i\left(\frac{x}{\lambda} - \nu t\right)\right] \tag{3.1.6}$$

representing waves traveling from left to right. Now we make use of the Planck formula ($E = h\nu$) and de Broglie formulas ($p = \frac{h}{\lambda}$) to replace ν and λ by their particle analogs. This gives

$$\Psi(x,t) = \exp\left[\frac{i(px - Et)}{\hbar}\right] \tag{3.1.7}$$

where

$$\hbar \equiv \frac{h}{2\pi} \tag{3.1.8}$$

Since Planck's constant occurs in most formulas with the denominator 2π, the $\hbar$ symbol was introduced by Paul Dirac. Equation 3.1.5 represents in some way the wavelike nature of a particle with energy E and momentum p. The time derivative of Equation 3.1.7 gives

$$\frac{\partial \Psi}{\partial t} = -\left(\frac{iE}{\hbar}\right)\exp\left[\frac{i(px - Et)}{\hbar}\right] \tag{3.1.9}$$

Thus from a simple comparison of Equations 3.1.7 and 3.1.9

$$i\hbar\frac{\partial \Psi}{\partial t} = E\Psi \tag{3.1.10}$$

or analogously differentiation of Equation 3.1.9 with respect to x

$$-i\hbar\frac{\partial \Psi}{\partial x} = p\Psi \tag{3.1.11}$$

and then the second derivative

$$-\hbar^2\frac{\partial^2 \Psi}{\partial x^2} = p^2\Psi \tag{3.1.12}$$

The energy and momentum for a *nonrelativistic* free particle (i.e., all energy is kinetic with no potential energy involved) are related by

$$E = \frac{1}{2}mv^2 = \frac{p^2}{2m} \tag{3.1.13}$$

Substituting Equations 3.1.12 and 3.1.10 into Equation 3.1.13 shows that $\Psi(x,t)$ satisfies the following partial differential equation

$$i\hbar\frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \Psi}{\partial x^2} \tag{3.1.14}$$

Equation 3.1.14 is the applicable differential equation describing the wavefunction of a *free particle* that is not bound by any external forces or equivalently not in a region where its potential energy $V(x,t)$ varies.

For a particle with a non-zero potential energy $V(x)$, the total energy E is then a sum of kinetics and potential energies

$$E = \frac{p^2}{2m} + V(x) \tag{3.1.15}$$

we postulate that Equation 3.1.3 for matter waves can be generalized to

$$\underbrace{i\hbar\frac{\partial \Psi(x,t)}{\partial t} = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\Psi(x,t)}_{\text{time-dependent Schrödinger equation in 1D}} \tag{3.1.16}$$

For matter waves in three dimensions, Equation 3.1.6 is then expanded

$$\underbrace{i\hbar\frac{\partial}{\partial t}\Psi(\vec{r},t)=\left[-\frac{\hbar^2}{2m}\nabla^2+V(\vec{r})\right]\Psi(\vec{r},t)}_{\text{time-dependent Schrödinger equation in 3D}} \tag{3.1.17}$$

Here the potential energy and the wavefunctions Ψ depend on the three space coordinates x, y, z, which we write for brevity as $\vec{r}$. Notice that the potential energy is assumed to depend on position only and not time (i.e., particle motion). This is applicable for conservative forces that a potential energy function $V(\vec{r})$ can be formulated.

The Laplacian Operator

The three second derivatives in parentheses together are called the Laplacian operator, or del-squared,

$$\begin{aligned}\nabla^2 &= \nabla\cdot\nabla \\ &= \left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\end{aligned} \tag{3.1.18}$$

with the del operator,

$$\nabla=\left(\vec{x}\frac{\partial}{\partial x}+\vec{y}\frac{\partial}{\partial y}+\vec{z}\frac{\partial}{\partial z}\right) \tag{3.1.19}$$

Remember from basic calculus that when the del operator is direclty operates on a field (e.g, $\nabla f(x,y,x)$, it denotes the gradient (i.e, the locally steepest slope) of the field. The symbols with arrows in Equation 3.1.19 are unit vectors.

Equation 3.1.17 is the *time-dependent Schrödinger* equation describing the wavefunction amplitude $\Psi(\vec{r},t)$ of matter waves associated with the particle within a specified potential $V(\vec{r})$. Its formulation in 1926 represents the start of modern quantum mechanics (Heisenberg in 1925 proposed another version known as matrix mechanics).

For *conservative* systems, the energy is a constant, and the time-dependent factor from Equation 3.1.7 can be separated from the space-only factor (via the *Separation of Variables* technique discussed in Section 2.2)

$$\Psi(\vec{r},t)=\psi(\vec{r})e^{-iEt/\hbar} \tag{3.1.20}$$

where $\psi(\vec{r})$ is a wavefunction dependent (or time-independent) wavefuction that only depends on space coordinates. Putting Equation 3.1.20 into Equation 3.1.17 and cancelling the exponential factors, we obtain the time-independent Schrödinger equation:

$$\underbrace{\left[-\frac{\hbar^2}{2m}\nabla^2+V(\vec{r})\right]\psi(\vec{r})=E\psi(\vec{r})}_{\text{time-independent Schrödinger equation}} \tag{3.1.21}$$

The overall form of the Equation 3.1.21 is not unusual or unexpected as it uses the principle of the **conservation of energy**. Most of our applications of quantum mechanics to chemistry will be based on this equation (with the exception of spectroscopy). The terms of the time-independent Schrödinger equation can then be interpreted as total energy of the system, equal to the system kinetic energy plus the system potential energy. In this respect, it is just the same as in classical physics.

Time Dependence to the Wavefunctions

Notice that the wavefunctions used with the time-independent Schrödinger equation (i.e., $\psi(\vec{r})$) do not have explicit t dependences like the wavefunctions of time-dependent analog in Equation 3.1.17 (i.e., $\Psi(\vec{r},t)$). That does **not imply** that there is no time dependence to the wavefunction. Equation 3.1.20 argues that the time-dependent (i.e., full spatial and temporal) wavefunction ($\Psi(\vec{r},t)$) differs from from the time-independent (i.e., spatial only) wavefunction $\psi(\vec{r})$ by a "phase factor" of constant magnitude. Using the Euler relationship in Equation 3.1.4, the total wavefunction above can be expanded

$$\Psi(\vec{r},t)=\psi(\vec{r})\left(\cos\frac{Et}{\hbar}-i\,\sin\frac{Et}{\hbar}\right) \tag{3.1.22}$$

This means the total wavefunction has a **complex behavior** with a real part and an imaginary part. Moreover, using the trigonometry identity $\sin(\theta)=\cos(\theta-\pi/2)$ Equation 3.1.22 can further simplified to

$$\Psi(\vec{r},t) = \psi(\vec{r})\cos\left(\frac{Et}{\hbar}\right) - i\psi(\vec{r})\cos\left(\frac{Et}{\hbar} - \frac{\pi}{2}\right)$$

This shows that both the real and the imaginary components of the total wavefunction oscillate, the imaginary part of the total wavefunction oscillates **out of phase** by $\frac{\pi}{2}$ with respect to the real part.

Note that while all wavefunctions have a time-dependence, that dependence may not impact in simple quantum problems as the next sections discuss and can often be ignored.

Before we embark on this, however, let us pause to comment on the validity of quantum mechanics. Despite its weirdness, its abstractness, and its strange view of the universe as a place of randomness and unpredictability, quantum theory has been subject to intense experimental scrutiny. It has been found to agree with experiments to better than $10^{-10}\%$ for all cases studied so far. When the Schrödinger Equation is combined with a quantum description of the electromagnetic field, a theory known as *quantum electrodynamics*, the result is one of the most accurate theories of matter that has ever been put forth. Keeping this in mind, let us forge ahead in our discussion of the quantum universe and how to apply quantum theory to both model and real situations.

3.1: The Schrödinger Equation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.1: The Schrödinger Equation** by Michael Fowler is licensed CC BY-NC-SA 4.0.

3.2: Linear Operators in Quantum Mechanics

Learning Objectives

- Classical-Mechanical quantities are represented by linear operators in Quantum Mechanics
- Understand that "algebra" of scalars and functions do not always to operators (specifically the commutative property)

The bracketed object in the time-independent Schrödinger Equation (in 1D)

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\right]\psi(\vec{r}) = E\psi(\vec{r}) \tag{3.2.1}$$

is called an *operator*. An operator is a generalization of the concept of a function applied to a function. Whereas a function is a rule for turning one number into another, an operator is a rule for turning one function into another. For the time-independent Schrödinger Equation, the operator of relevance is the Hamiltonian operator (often just called the Hamiltonian) and is the most ubiquitous operator in quantum mechanics.

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})$$

We often (but not always) indicate that an object is an operator by placing a 'hat' over it, eg, $\hat{H}$. So time-independent Schrödinger Equation can then be simplified from Equation 3.2.1 to

$$\hat{H}\psi(\vec{r}) = E\psi(\vec{r}) \tag{3.2.2}$$

Equation 3.2.2 says that the Hamiltonian operator operates on the wavefunction to produce the energy, which is a scalar (i.e., a number, a quantity and observable) times the wavefunction. Such an equation, where the operator, operating on a function, produces a constant times the function, is called an **eigenvalue equation**. The function is called an eigenfunction, and the resulting numerical value is called the eigenvalue. Eigen here is the German word meaning self or own. We will discuss this in detail in later Sections.

Fundamental Properties of Operators

Most properties of operators are straightforward, but they are summarized below for completeness.

1. The sum and difference of two operators $\hat{A}$ and $\hat{B}$ are given by

$$(\hat{A} \pm \hat{B})f = \hat{A}f \pm \hat{B}f$$

2. The product of two operators is defined by

$$\hat{A}\hat{B}f \equiv \hat{A}[\hat{B}f]$$

3. Two operators are equal if

$$\hat{A}f = \hat{B}f$$

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

$$\hat{1}f = f$$

5. The n-th power of an operator $\hat{A}^n$ is defined as n successive applications of the operator, e.g.

$$\hat{A}^2 f = \hat{A}\hat{A}f$$

6. The *associative law* holds for operators

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

7. The *commutative law* does **not** generally hold for operators. In general, but not always,

$$\hat{A}\hat{B} \neq \hat{B}\hat{A}. \tag{3.2.3}$$

To help identify if the inequality in Equation 3.2.3 holds for any two specific operators, we define the commutator.

Definition: The Commutator

It is convenient to define the **commutator** of $\hat{A}$ and $\hat{B}$

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$$

If $\hat{A}$ and $\hat{B}$ commute, then

$$[\hat{A}, \hat{B}] = 0.$$

If the commutator is not zero, the order of operating matters and the operators are said to "not commute." Moreover, this property applies

$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}].$$

Linear Operators

The action of an operator that turns the function $f(x)$ into the function $g(x)$ is represented by

$$\hat{A}f(x) = g(x) \tag{3.2.4}$$

The most common kind of operator encountered are *linear operators* which satisfies the following two conditions:

$$\underset{\text{Condition A}}{\hat{O}(f(x) + g(x)) = \hat{O}f(x) + \hat{O}g(x)} \tag{3.2.5}$$

and

$$\underset{\text{Condition B}}{\hat{O}cf(x) = c\hat{O}f(x)} \tag{3.2.6}$$

where

- $\hat{O}$ is a linear operator,
- c is a constant that can be a complex number ($c = a + ib$), and
- $f(x)$ and $g(x)$ are functions of x

If an operator fails to satisfy either Equations 3.2.5 or 3.2.6 then it is not a linear operator.

Example 3.2.1

Is this operator $\hat{O} = -i\hbar\frac{d}{dx}$ linear?

Solution

To confirm is an operator is linear, both conditions in Equation 3.2.6 must be demonstrated.

Condition A (Equation 3.2.5):

$$\hat{O}(f(x) + g(x)) = -i\hbar\frac{d}{dx}(f(x) + g(x))$$

From basic calculus, we know that we can use the **sum rule for differentiation**

$$\begin{aligned}\hat{O}(f(x) + g(x)) &= -i\hbar\frac{d}{dx}f(x) - i\hbar\frac{d}{dx}g(x) \\ &= \hat{O}f(x) + \hat{O}g(x) \quad \checkmark\end{aligned}$$

Condition A is confirmed. Does Condition B (Equation 3.2.6) hold?

$$\hat{O}cf(x) = -i\hbar\frac{d}{dx}cf(x)$$

Also from basic calculus, this can be factored out of the derivative

$$\hat{O}cf(x) = -ci\hbar\frac{d}{dx}f(x)$$
$$= c\hat{O}f(x) \quad \checkmark$$

Yes. This operator is a linear operator (this is the linear momentum operator).

? Exercise 3.2.1

Confirm if the square root operator $\sqrt{f(x)}$ linear or not?

The most operators encountered in quantum mechanics are *linear operators*.

Hermitian Operators

An important property of operators is suggested by considering the Hamiltonian for the particle in a box:

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$$

Let $f(x)$ and $g(x)$ be arbitrary functions which obey the same boundary values as the eigenfunctions of $\hat{H}$ (e.g., they vanish at $x = 0$ and $x = a$). Consider the integral

$$\int_0^a f(x)\,\hat{H}\,g(x)\,\mathrm{d}x = -\frac{\hbar^2}{2m}\int_0^a f(x)\,g''(x)\,\mathrm{d}x$$

Now, using integration by parts,

$$\int_0^a f(x)\,g''(x)\,\mathrm{d}x = -\int_0^a f'(x)\,g'(x)\,\mathrm{d}x + f(x)\,g'(x)\Big|_0^a \tag{3.2.7}$$

The boundary terms vanish by the assumed conditions on f and g. A second integration by parts transforms Equation 3.2.7 to

$$\int_0^a f''(x)\,g(x)\,\mathrm{d}x - f'(x)\,g(x)\Big|_0^a$$

It follows therefore that

$$\int_0^a f(x)\,\hat{H}\,g(x)\,\mathrm{d}x = \int_0^a g(x)\,\hat{H}\,f(x)\,\mathrm{d}x$$

An obvious generalization for complex functions will read

$$\int_0^a f^*(x)\,\hat{H}\,g(x)\,\mathrm{d}x = \left(\int_0^a g^*(x)\,\hat{H}\,f(x)\,\mathrm{d}x\right)^*$$

In mathematical terminology, an operator $\hat{A}$ for which

$$\int f^*\,\hat{A}\,g\,\mathrm{d}\tau = \left(\int g^*\,\hat{A}\,f\,\mathrm{d}\tau\right)^*$$

for all functions f and g which obey specified boundary conditions is classified as *Hermitian* or *self-adjoint*. Evidently, the Hamiltonian is a **Hermitian operator**. It is postulated that *all* quantum-mechanical operators that represent dynamical variables are Hermitian. The term is also used for specific times of matrices in linear algebra courses.

> *All quantum-mechanical operators that represent dynamical variables are Hermitian.*

3.2: Linear Operators in Quantum Mechanics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.2: Linear Operators in Quantum Mechanics** by S. M. Blinder is licensed CC BY-NC-SA 4.0.

3.3: The Schrödinger Equation is an Eigenvalue Problem

Learning Objectives

- To recognize that each quantum mechanical observable is determined by solve by an eigenvalue problem with different operators for different observable
- Confirm if a specific wavefunction is an eigenfunction of a specific operation and extract the corresponding obserable (the eigenvalue)
- To recognize that the Schrödinger equation, just like all measurable, is also an eigenvalue problem with the eigenvalue ascribed to total energy
- Identity and manipulate several common quantum mechanical operators

As per the definition, an operator acting on a function gives another function, however a special case occurs when the generated function is proportional to the original

$$\hat{A}\psi \propto \psi \tag{3.3.2}$$

This case can be expressed in terms of a equality by introducing a proportionality constant k

$$\hat{A}\psi = k\psi \tag{3.3.3}$$

Not all functions will solve an equation like in Equation 3.3.3. If a function does, then ψ is known as an **eigenfunction** and the constant k is called its **eigenvalue** (these terms are hybrids with German, the purely English equivalents being "characteristic function" and "characteristic value", respectively). Solving eigenvalue problems are discussed in most linear algebra courses.

In quantum mechanics, every experimental measurable a is the eigenvalue of a specific operator ($\hat{A}$):

$$\hat{A}\psi = a\psi \tag{3.3.4}$$

The a eigenvalues represents the possible measured values of the $\hat{A}$ operator. Classically, a would be allowed to vary continuously, but in quantum mechanics, a typically has only a sub-set of allowed values (hence the quantum aspect). Both time-dependent and time-independent Schrödinger equations are the best known instances of an eigenvalue equations in quantum mechanics, with its eigenvalues corresponding to the allowed energy levels of the quantum system.

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\right]\psi(\vec{r}) = E\psi(\vec{r}) \tag{3.3.5}$$

The object on the left that acts on $\psi(x)$ is an example of an **operator**.

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\right] \tag{3.3.6}$$

In effect, what is says to do is "take the second derivative of $\psi(x)$, multiply the result by $-(\hbar^2/2m)$ and then add $V(x)\psi(x)$ to the result of that." Quantum mechanics involves many different types of operators. This one, however, plays a special role because it appears on the left side of the Schrödinger equation. It is called the **Hamiltonian** *operator* and is denoted as

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}) \tag{3.3.7}$$

Therefore, the time-dependent Schrödinger equation can be (and it more commonly) written as

$$\hat{H}\psi(x,t) = i\hbar\frac{\partial}{\partial t}\psi(x,t) \tag{3.3.8}$$

and the time-independent Schrödinger equation

$$\hat{H}\psi(x) = E\psi(x) \tag{3.3.9}$$

Note that the functional form of Equation 3.3.9 is the same as the general eigenvalue equation in Equation 3.3.3 where the eigenvalues are the (allowed) total energies (E).

The Hamiltonian, named after the Irish mathematician Hamilton, comes from the formulation of Classical Mechanics that is based on the total energy, $H = T + V$, rather than Newton's second law, $F = ma$. Equation 3.3.9 says that the Hamiltonian operator operates on the wavefunction to produce the energy E, which is a scalar (e.g., expressed in Joules) times the wavefunction.

Correspondence Principle

Note that $\hat{H}$ is derived from the classical energy $p^2/2m + V(x)$ simply by replacing $p \rightarrow -i\hbar(d/dx)$. This is an example of the **Correspondence Principle** initially proposed by Niels Bohr that states that the behavior of systems described by quantum theory reproduces classical physics in the limit of large quantum numbers.

It is a general principle of Quantum Mechanics that there is an operator for every physical observable. A physical observable is anything that can be measured. If the wavefunction that describes a system is an eigenfunction of an operator, then the value of the associated observable is extracted from the eigenfunction by operating on the eigenfunction with the appropriate operator. The value of the observable for the system is then the eigenvalue, and the system is said to be in an eigenstate. Equation 3.3.9 states this principle mathematically for the case of energy as the observable. If the wavefunction is not the eigenfunction of the operation, then the measurement will give an eigenvalue (by definition), but not necessarily the same one for each measurement (this will be discussed in more detail in later section).

Common Operators

Although we could theoretically come up with an infinite number of operators, in practice there are a few which are much more important than any others.

- **Linear Momentum:**

  The linear momentum operator of a particle moving in one dimension (the x-direction) is

  $$\hat{p}_x = -i\hbar\frac{\partial}{\partial x} \tag{3.3.10}$$

  and can be generalized in three dimensions:

  $$\hat{\vec{p}} = -i\hbar\nabla \tag{3.3.11}$$

- **Position**

  The position operator of a particle moving in one dimension (the x-direction) is

  $$\hat{x} = x \tag{3.3.12}$$

  and can be generalized in three dimensions:

  $$\hat{\vec{r}} = \vec{r} \tag{3.3.13}$$

  where $\vec{r} = (x, y, z)$.

- **Kinetic Energy**

  Classically, the kinetic energy of a particle moving in one dimension (the x-direction), in terms of momentum, is

  $$KE_{classical} = \frac{p_x^2}{2m} \tag{3.3.14}$$

  Quantum mechanically, the corresponding kinetic energy operator is

  $$\hat{KE}_{quantum} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} \tag{3.3.15}$$

  and can be generalized in three dimensions:

$$\hat{KE}_{quantum} = -\frac{\hbar^2}{2m}\nabla^2 \tag{3.3.16}$$

- **Angular Momentum:**

  Angular momentum requires a more complex discussion, but is the cross product of the position operator $\hat{\vec{r}}$ and the momentum operator $\hat{p}$

$$\hat{\vec{L}} = -i\hbar(\vec{r} \times \nabla) \tag{3.3.17}$$

- **Hamiltonian:**

  The Hamiltonian operator corresponds to the total energy of the system

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \tag{3.3.18}$$

  and it represents the total energy of the particle of mass m in the potential $V(x)$. The Hamiltonian in three dimensions is

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}) \tag{3.3.19}$$

- **Total Energy:**

  The energy operator from the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \hat{H}\Psi(x,t) \tag{3.3.20}$$

  The right hand side of Equation 3.3.7 is the Hamiltonian Operator. In addition determining system energies, the Hamiltonian operator dictates the time evolution of the wavefunction

$$\hat{H}\Psi(x,t) = i\hbar\frac{\partial\Psi(x,t)}{\partial t} \tag{3.3.21}$$

  This aspect will be discussed in more detail elsewhere.

Eigenstate, Eigenvalues, Wavefunctions, Measurables, and Observables

In general, the **wavefunction** gives the "state of the system" for the system under discussion. It stores all the information available to the observer about the system. Often in discussions of quantum mechanics, the terms *eigenstate* and *wavefunction* are used interchangeably. The term eigenvalue is used to designate the value of measurable quantity associated with the wavefunction.

- If you want to measure the energy of a particle, you have to operate on the wavefunction with the Hamiltonian operator (Equation 3.3.7).
- If you want to measure the momentum of a particle, you have to operate on wavefunction with the momentum operator (Equation 3.3.10).
- If you want to measure the position of a particle, you have to operate on wavefunction with the position operator (Equation 3.3.12).
- If you want to measure the kinetic energy of a particle, you have to operate on wavefunction with the kinetic energy operator (Equation 3.3.15).

When discussing the eigenstates of the Hamiltonian ($\hat{H}$), the associated eigenvalues represent energies and within the context of the momentum operators, the associated eigenvalues refer to the momentum of the particle. However, not all wavefunctions (ψ) are eigenstates of an operator (ϕ) – and if they are not, they can be usually be written as superpositions of eigenstates.

$$\psi = \sum_i c_i\phi_i$$

This will be discussed in more detail in later sections.

While the wavefunction may not be the eigenstate of an observable, when that operator operates on that wavefunction, the wavefunction **becomes** an eigenstate of that observable and only eigenvalues can be observed. Another way to say this is that the

wavefunction "collapses" into an eigenstate of the observable. Because quantum mechanical operators have different forms, their associated eigenstates are similarly often (i.e., most of the time) different. For example, when a wavefunction is an eigenstate of total energy, it will **not** be an eigenstate of momentum.

> *If a wavefunction is an eigenstate of one operator, (e.g., momentum), that state is not necessarily an eigenstate of a different operator (e.g., energy), although not always.*

The wavefunction immediately after a measurement is an eigenstate of the operator associated with this measurement. What happens to the wavefunction after the measurement is a different topic.

Example 3.3.1

Confirm that the following wavefunctions are eigenstates of linear momentum and kinetic energy (or neither or both):

a. $\psi = A\sin(ax)$
b. $\psi = Ne^{-ix/\hbar}$

Strategy

This question is asking if the eigenvalue equation holds for the operators and these wavefunctions. This is just asking if these wavefunctions are solutions to Equation 3.3.3 using the operators in Equations 3.3.10 and 3.3.15, i.e., are these equations true:

$$\hat{p}_x\psi = p_x\psi \tag{3.3.22}$$

$$\hat{KE}\psi = KE\psi \tag{3.3.23}$$

where p_x and KE are the measurables (eigenvalues) for these operators.

Solution a

Let's evaluate the left side of the linear momentum eigenvalue problem (Equation 3.3.22)

$$-i\hbar\frac{\partial}{\partial x}A\sin(ax) = -i\hbar Aa\cos(ax)$$

and compare to the the right side of Equation 3.3.22

$$p_x A\sin(ax)$$

These are **not** the same so this wavefunction is not an eigenstate of momentum.

Let's look at the left side of the kinetic energy eigenvalue problem (Equation 3.3.23)

$$\begin{aligned}-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}A\sin(ax) &= -\frac{\hbar^2}{2m}\frac{\partial}{\partial x}Aa\cos(ax)\\ &= +\frac{\hbar^2}{2m}Aa^2\sin(ax)\end{aligned}$$

and compare to the right side

$$KEA\sin(ax)$$

These are same, so this specific wavefunction is an eigenstate of kinetic energy. Moreover, the measured kinetic energy will be

$$KE = +\frac{\hbar^2}{2m}a^2$$

Solution b

Let's look at the left side of Equation 3.3.22 for linear momentum

$$-i\hbar\frac{\partial}{\partial x}Ne^{-ix/\hbar} = -Ne^{-ix/\hbar}$$

and the right side of Equation 3.3.22

$$p_x N e^{-ix/\hbar}$$

These are the same so this wavefunction is an eigenstate of momentum with momentum $p_x = -N$.

Let's look at the left side of Equation 3.3.23 for kinetic energy

$$\begin{aligned} -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} N e^{-ix/\hbar} &= +i\frac{\hbar}{2m}\frac{\partial}{\partial x} N e^{-ix/\hbar} \\ &= +\frac{1}{2m} N e^{-ix/\hbar} \end{aligned}$$

and the right side

$$KE N e^{-ix/\hbar}$$

These are same so this wavefunction is an eigenstate of kinetic energy. And the measured kinetic energy will be

$$KE = \frac{1}{2m}$$

This wavefunction is an eigenstate of both momentum and kinetic energy.

Exercise 3.3.1

Are $\psi = Me^{-bx}$ functions eigenstates of linear momentum and kinetic energy (or neither or both)?

3.3: The Schrödinger Equation is an Eigenvalue Problem is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.3: The Schrödinger Equation is an Eigenvalue Problem** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

3.4: The Quantum Mechanical Free Particle

The simplest system in quantum mechanics has the potential energy $V = 0$ everywhere. This is called a *free particle* since it has no forces acting on it. We consider the one-dimensional case, with motion only in the x-direction, giving the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} = E\psi(x) \tag{3.4.1}$$

or more simply

$$\hat{H}\psi(x) = E\psi(x) \tag{3.4.2}$$

with

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \tag{3.4.3}$$

Total derivatives can be used since there is one independent variable. The equation simplifies to

$$\psi''(x) + k^2\psi(x) = 0 \tag{3.4.4}$$

with the definition

$$k^2 \equiv \frac{2mE}{\hbar^2} \tag{3.4.5}$$

Possible solutions of Equation 3.4.4 are

$$\psi(x) = const\begin{Bmatrix}\sin(kx)\\ \cos(kx)\\ e^{\pm ikx}\end{Bmatrix} \tag{3.4.6}$$

There is no restriction on the value of k. Thus a free particle, even in quantum mechanics, can have any non-negative value of the energy

$$E = \frac{\hbar^2k^2}{2m} \geq 0 \tag{3.4.7}$$

The energy levels in this case are *not* quantized and correspond to the same continuum of kinetic energy shown by a classical particle.

Example 3.4.1: Eigenstates of the Hamiltonian of a free particle

Consider for a free particle (i.e., $V(x) = 0$) moving in one dimension.

a. Demonstrate that $\sin(kx)$, $\cos(kx)$, and $e^{\pm ikx}$ are indeed eigenfunctions of $\hat{H}$.
b. What are their corresponding eigenstates?
c. What are the energies not quantized?

Exercise 3.4.1

Consider for a free particle (i.e., $V(x) = 0$) moving in one dimension.

a. Demonstrate that $\cos(kx)$ and $e^{\pm ikx}$ are eigenfunctions of $\hat{H}$.
b. What are their corresponding eigenstates?
c. Is momentum quantized for a free particle?

It is of interest also to consider the x-component of linear momentum for the free-particle solutions (Equation 3.4.6). The eigenvalue equation for momentum is

$$\hat{p}_x\psi(x) = -i\hbar\frac{d\psi(x)}{dx} = p\psi(x) \tag{3.4.14}$$

where we have denoted the momentum eigenvalue as p.

It is easily shown that neither of the functions $\sin kx$ or $\cos kx$ from Equation 3.4.6 is an eigenfunction of $\hat{p}_x$. However, $e^{\pm ikx}$ are both eigenfunctions with eigenvalues $p = \pm\hbar k$, respectively. Evidently the momentum p can take on any real value between -∞ and +∞. The kinetic energy, equal to $E = p^2/2m$, can correspondingly have any value between 0 and +∞.

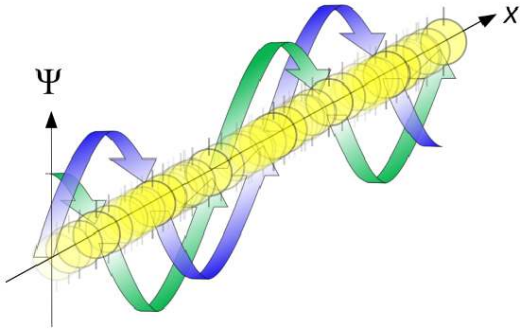


Figure 3.4.1: Propagation of quantum free particle (i.e. de Broglie waves) in 1D space. The real part of the complex amplitude is blue and imaginary part is green. The probability (shown as the yellow opacity) of finding the particle at a given point x is spread out evenly over space, there is no definite position of the particle. Image used with permission (Public Domain; Maschen).

> *A free particle, even in quantum mechanics, can have any non-negative value of the energy*

Example 3.4.2: Eigenstates of Momentum of a Free Particle

Consider for a free particle (i.e., $V(x) = 0$) moving in one dimension.

a. Demonstrate that $e^{\pm ikx}$ are eigenfunctions of $\hat{p}_x$.
b. What are their corresponding eigenstates?

Exercise 3.4.2

Show that neither $\sin kx$ nor $\cos kx$ are eigenstates of $\hat{p}_x$ for a free particle (i.e., $V(x) = 0$) moving in one dimension.

Superpositions

The functions $\sin kx$ and $\cos kx$, while not eigenfunctions of $\hat{p}_x$, are each *superpositions* of the two eigenfunctions $e^{\pm ikx}$, by virtue of the trigonometric identities

$$\cos(kx) = \frac{1}{2}(e^{ikx} + e^{-ikx}) \tag{3.4.15}$$

and

$$\sin(kx) = \frac{1}{2i}(e^{ikx} - e^{-ikx}) \tag{3.4.16}$$

The eigenfunction e^{ikx} for $k > 0$ represents the particle moving from left to right on the *x-axis* (blue curve in Figure 3.4.2) with momentum $p > 0$. Correspondingly, e^{-ikx} represents motion from right to left with $p < 0$ (red curve in Figure 3.4.2).

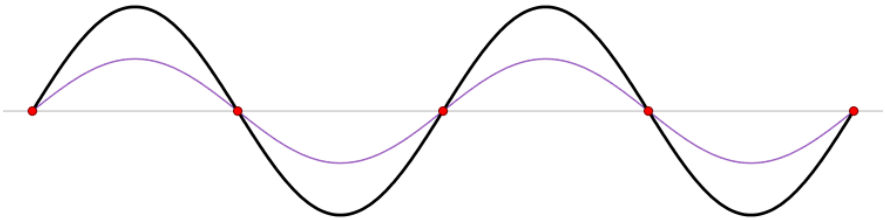


Figure 3.4.2: Two waves traveling in opposite directions across the same medium combine linearly. In this animation, both waves have the same wavelength and the sum of amplitudes results in a standing wave. Image used with permission (Public Domain; LucasVB)

The functions $\sin(kx)$ and $\cos(kx)$ represent *standing waves*, obtained by superposition of opposing complex waves. Although these two trigonometry functions are not eigenfunctions of $\hat{p}_x$, they *are* eigenfunctions of $\hat{p}_x^2$ and hence they are eigenfunctions of the Hamiltonian operator $\hat{H}$.

Not every wavefunction is an eigenstate

It is important to note that while $\sin kx$ nor $\cos kx$ are **not** eigenstates of momentum, these functions can be decomposed into a combination of functions that are eigenstates (Equation 3.4.14 and 3.4.16). The ability to decompose wavefunctions into (linear) combinations of eigenstates is used throughout quantum mechanics and will be discussed in more detail in further Sections.

Wavepackets - The Superposition Principle in Action

Consider the wavefunction corresponding to a free electron traveling through a cathode ray tube. The electron leaves the cathode, shoots through the vacuum, and impinges on the screen. At an intermediate point in this process, it is moving through the vacuum and the wavefunction must be nonzero over some volume, but zero in the places the electron has not possibly reached yet, and zero in the places it has definitely left.

However, if the electron has a precise energy (which is all kinetic energy), it also has a precise *momentum*. De Broglie argued that this necessarily implies that the electron wave has a precise *wavelength*. However, since it is a free particle ($V = 0$, the only wave with a *precise* wavelength λ has the form

$$\psi(x,t) = Ae^{i(kx-\omega t)} \tag{3.4.17}$$

where $k = \frac{2\pi}{\lambda}$ and $\omega = 2\pi\nu$. The wavefunctions of Equation 3.4.17 are the solutions to the full time-independent Schrödinger equation (not just the spatial component discussed above).

The problem with using the wavefunctions of Equation 3.4.17 to represent free particles is that plane sine waves extend to infinity in both spatial directions, so cannot represent a particle whose wavefunction is nonzero in a limited region of space. Physically, that it must have a wavefunction that goes to zero far away in either direction. A localized wavefunction of this type is called a "**wavepacket**". We shall discover that a wavepacket can be constructed by adding plane waves together. To represent a localized particle, we must superpose waves having different wavelengths. Now, the plane waves we add together will individually be solutions of the Schrödinger equation.

The Schrödinger equation is a linear Differential equation

As discussed in previously for a classical wave equation, the quantum equivalent, i.e., Schrödinger equation (both time dependent and independent) are *linear* equations. That is, if $\psi_1(x,y,z,t), \psi_2(x,y,z,t)$ are both solutions of the equation, then so is

$$\psi(x,y,z,t) = c_1\psi_1(x,y,z,t) + c_2\psi_2(x,y,z,t) \tag{3.4.18}$$

where c_1 and c_2 are arbitrary constants, as is easy to check. This is called the *Principle of Superposition*.

The essential point is that in Schrödinger's equation every term contains a factor ψ, but no term contains a factor ψ^2 (or a higher power). That is what is meant by a "linear" equation. If the equation *did* contain a constant term, or a term including ψ^2, superposition wouldn't work—the sum of two solutions to the equation would not itself be a solution to the equation.

In fact, we have been assuming this linearity all along: when we analyze interference and diffraction of waves, we just add the two wave amplitudes at each spot. For the double slit, we take it that if the wave radiating from one slit satisfies the wave equation, then adding the two waves together will give a new wave which also satisfies the equation.

Building a Wavepacket

If we add together two sine waves with frequencies close together, we get "beats." This pattern can be viewed as a string of wavepackets. For example considering adding to sine waves together with spatial and temporal frequencies of ($k \pm \Delta k$) and ($\omega \pm \Delta\omega$), respectively.

$$\sin\left[(k-\Delta k)x-(\omega-\Delta\omega)t\right]+\sin\left[(k+\Delta k)x-(\omega+\Delta\omega)t\right]) \tag{3.4.19}$$

If we use the trigonometric addition formula, the sum in Equation 3.4.19 can be simplified to

$$2\sin(kx-\omega t)\cos\left[(\Delta k)x-(\Delta\omega)t\right] \tag{3.4.20}$$

Equation 3.4.20 exhibits the phenomenon of beats between waves close in frequency. The first term, $\sin(kx-\omega t)$, oscillates at the average of the two frequencies. It is modulated by the slowly varying second term, often called the "**envelope function**", which oscillates once over a spatial extent of order $\pi/\Delta k$. This is the distance over which waves initially in phase at the origin become completely out of phase. Of course, going a further distance of order $\pi/\Delta k$, the waves will become synchronized again.

That is, beating two close frequencies together breaks up the continuous wave into a series of packets, the beats. To describe a single electron moving through space, we need a single packet. This can be achieved by superposing waves having a continuous distribution of wavelengths, or wave numbers within of order Δk, say, of k. In this case, the waves will be out of phase after a distance of order $\pi/\Delta k$, but since they have many different wavelengths, they will never get back in phase again.

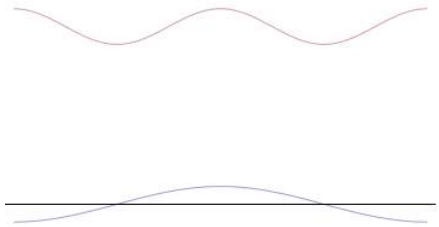


Figure 3.4.3: The superposition of several plane waves to form a wave packet. This wave packet becomes increasingly localized with the addition of many waves. The Fourier transform is a mathematical operation that separates a wave packet into its individual plane waves. Note that the waves shown here are real for illustrative purposes only, whereas in quantum mechanics the wavefunction is generally complex. Image used with permission (Public Domain; Teply).

We habe seen how two sine waves of equal amplitude close together in frequency produce beats: if the waves are in phase at the origin, as we go along the *x*-axis they gradually fall out of phase, and cancel each other at a distance $x = \frac{\pi}{2\Delta}$, where 2Δ is the difference in k of the two $\sin kx$ waves (Figure 3.4.3). (For the moment, we are ignoring the time development of these waves: we're just looking at *t* = 0.). If we continue along the *x*-axis to $\frac{\pi}{\Delta}$, the two waves will be back in phase again, this is the next beat. Now, if instead of adding *two* waves, we add *many* waves, all of different k, but with the k's taken from some small interval of size of order Δk, and all these waves are in phase at the origin, then, again, they will all stay more or less in phase for a distance of order $x = \frac{\pi}{2\Delta}$.

However, as we proceed past that point, the chances of them all getting back in phase again get rapidly smaller as we increase the number of different waves. This wavepacket perspective of a free particle is shown in Figure 3.4.4, where the wavefunction amplitude (both real and imaginary) is localized and tapers off on either side (contrast this image with Figure 3.4.1 that exhibits uniform amplitude over all space since it is an oscillatory eigenfunction.

Gaussian Wavepackets

Fortunately, there is a simple explicit mathematical realization of the addition of plane waves to form a localized function: the *Gaussian wavepacket:*

$$\psi(x,t=0)=Ae^{ik_0x}e^{-x^2/2\Delta^2} \tag{3.4.21}$$

where k_0 is an average wavenumber. Topically, the equation for Gaussian wavepackets resemble the equation for beating (Equation 3.4.20), but with a Gaussian *envelope function*. For this wavepacket to represent one electron, with the probability of finding the electron outside in a small section of length dx at x is negligible.



A time-dependent (moving) gaussian wavepacket (real or imaginary part). Image used with permission (Public Domain; Wikipedia)

Contributors

- Seymour Blinder (Professor Emeritus of Chemistry and Physics at the University of Michigan, Ann Arbor)
- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

3.4: The Quantum Mechanical Free Particle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.5: Wavefunctions Have a Probabilistic Interpretation

Learning Objectives

- To understand that wavefunctions can have probabilistic interpretations.
- To calculate the probabilities directly from a wavefunctions

For a single-particle system, the wavefunction $\Psi(\vec{r},t)$, or $\psi(\vec{r})$ for the time-independent case, represents the amplitude of the still vaguely defined matter waves. Since wavefunctions can in general be complex functions, the physical significance cannot be found from the function itself because the $\sqrt{-1}$ is not a property of the physical world. Rather, the physical significance is found in the product of the wavefunction and its complex conjugate, i.e. the absolute square of the wavefunction, which also is called the **square of the modulus** (also called absolute value).

$$P(\vec{r},t) = \Psi^*(\vec{r},t)\Psi(\vec{r},t) \tag{3.5.1}$$

$$= |\Psi(\vec{r},t)|^2 \tag{3.5.2}$$

where $\vec{r}$ is a vector (x,y,z) specifying a point in three-dimensional space. The square is used, rather than the modulus itself, just like the intensity of a light wave depends on the square of the electric field.

Born proposed in 1926, the most commonly accepted interpretation of the wavefunction **that the square of the modulus** (Equation 3.5.2) is **proportional** to the **probability density** (probability per unit volume) that the electron is in the volume $d\tau$ located at r_i. Since the wavefunction represents the wave properties of matter, the probability amplitude $P(x,t)$ will also exhibit wave-like behavior. Probability density is the three-dimensional analog of the diffraction pattern that appears on the two-dimensional screen in the double-slit diffraction experiment for electrons. The idea that we can understand the world of atoms and molecules only in terms of probabilities is disturbing to some, who are seeking more satisfying descriptions through ongoing research.

> *The Born interpretation therefore calls the wavefunction the probability amplitude, the absolute square of the wavefunction is called the probability density, and the probability density times a volume element in three-dimensional space ($d\tau$) is the probability P*

The probability that a single quantum particle moving in one spatial dimension will be found in a region $x \in [a,b]$ if a measurement of its location is performed is

$$P(x \in [a,b]) = \int_a^b |\psi(x)|^2 dx \tag{3.5.3}$$

In three dimensions, Equation 3.5.3 is represented differently

$$P(x \in [a,b]) = \int_V |\psi(\vec{r})|^2 d\tau \tag{3.5.4}$$

This integration extends over a specified volume (V) with the symbol $d\tau$ designating the appropriate volume element (including a Jacobian) of the coordinate system adopted:

- Cartesian:

$$d\tau = dx\,dy\,dz$$

- Spherical:

$$d\tau = r^2 \sin\phi\, dr\, d\theta\, d\phi$$

- Cylindrical:

$$d\tau = r\, dr\, d\phi\, dz.$$

For rectilinear Cartesian space, Equation 3.5.4 can be is expanded with dimension explicitly indicated

$$P(x \in [a,b]) = \int_{a_x}^{b_x} \int_{a_y}^{b_y} \int_{a_z}^{b_z} |\psi(x,y,z)|^2 dx\, dy\, dz \tag{3.5.5}$$

where the limits of integration are selected to encompass the volume V of consideration.

The Born interpretation (Equation 3.5.2) of relating the wavefunction to probability forces certain demands on its mathematical behavior of wavefunctions and not any mathematical function can be a valid wavefunction.

Required Properties of Wavefunction

- The wavefunction must be a *single-valued function* of all its coordinates, since the probability density ought to be *uniquely determined* at each point in space.
- The wavefunction should be both *finite* as an infinite probability has no meaning.
- The wavefunction should be *continuous everywhere*, as expected for a physically-meaningful probability density.

The conditions that the wavefunction be single-valued, finite and continuous--in short, "well behaved"-- lead to restrictions on solutions of the Schrödinger equation such that only certain values of the energy and other dynamical variables are allowed. This is called *quantization* and is in the feature that gives *quantum* mechanics its name.

It is important to note that this interpretation implies the wavefunction does **not** mean the particle is distributed over a large region as a sort of "charge cloud". The wavefunction is used to describe the electron motion that behaves like waves and satisfies a wave equation. This is akin to how a grade distribution in a large class does not represent a smearing of grades for a single student, but only makes sense when taking into account that the distribution is the result of many measurables (e.g., student performances).

Example 3.4.1

Show that the square of the modulus of $\Psi(\vec{r},t) = \psi(\vec{r})e^{-i\omega t}$ is time independent. What insight regarding stationary states do you gain from this proof?

Example 3.4.2

According to the Born interpretation, what is the physical significance of $e\psi^*(r_0)(r_0)d\tau$?

3.5: Wavefunctions Have a Probabilistic Interpretation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.4: Wavefunctions Have a Probabilistic Interpretation** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

3.6: The Energy of a Particle in a Box is Quantized

Learning Objectives

- Solve the particle-in-a-box model used to describing a trapped particle in 1 D well
- Characterize the particle-in-a-box eigenstates (i.e., wavefunctions) and the eigenenergies as a function of the quantum number
- Demonstrate that the eigenstates are orthogonal

The *particle in the box* model system is the simplest non-trivial application of the Schrödinger equation, but one which illustrates many of the fundamental concepts of quantum mechanics. For a particle moving in one dimension (again along the x- axis), the Schrödinger equation can be written

$$-\frac{\hbar^2}{2m}\psi''(x) + V(x)\psi(x) = E\psi(x)$$

Assume that the particle can move freely between two endpoints $x = 0$ and $x = L$, but cannot penetrate past either end. This is equivalent to a potential energy dependent on x with

$$V(x) = \begin{cases} 0 & 0 \le x \le L \\ \infty & x < 0 \; and \; x > L \end{cases}$$

This potential is represented in Figure 3.5.1 . The infinite potential energy constitutes an impenetrable barrier since the particle would have an infinite potential energy if found there, which is clearly impossible.

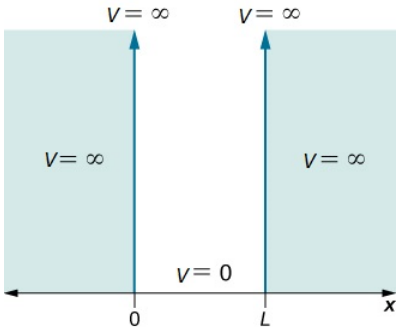


Figure 3.5.1 : The barriers outside a one-dimensional box have infinitely large potential, while the interior of the box has a constant, zero potential. (CC-BY 4.0; OpenStax).

The particle is thus bound to a *"potential well"* since the particle cannot penetrate beyond $x = 0$ or $x = L$

$$\psi(x) = 0 \quad for \;\; x < 0 \;\; and \;\; x > L \tag{3.6.1}$$

By the requirement that the wavefunction be continuous, it must be true as well that

$$\psi(0) = 0 \quad and \quad \psi(L) = 0 \tag{3.6.2}$$

which constitutes a pair of boundary conditions on the wavefunction within the box. Inside the box, $V(x) = 0$, so the Schrödinger equation reduces to the free-particle form:

$$-\frac{\hbar^2}{2m}\psi''(x) = E\psi(x) \tag{3.6.3}$$

with $0 \le x \le L$.

We again have the differential equation

$$\psi''(x) + k^2\psi(x) = 0 \tag{3.6.4}$$

with

$$k^2 = \frac{2mE}{\hbar^2} \tag{3.6.5}$$

The general solution can be written

$$\psi(x) = A \sin\ kx + B \cos\ kx \tag{3.6.6}$$

where A and B are constants to be determined by the boundary conditions in Equation 3.6.2. By the first condition, we find

$$\psi(0) = A \sin 0 + B \cos 0 = B = 0 \tag{3.6.7}$$

The second boundary condition at $x = L$ then implies

$$\psi(a) = A \sin kL = 0 \tag{3.6.8}$$

It is assumed that $A \neq 0$, for otherwise $\psi(x)$ would be zero everywhere and the particle would disappear (i.e., the trivial solution). The condition that $\sin kx = 0$ implies that

$$kL = n\pi \tag{3.6.9}$$

where n is a integer, positive, negative or zero. The case $n = 0$ must be excluded, for then $k = 0$ and again $\psi(x)$ would vanish everywhere. Eliminating k between Equation 3.6.4 and 3.6.9, we obtain

$$E_n = \frac{\hbar^2\pi^2}{2mL^2} n^2 = \frac{h^2}{8mL^2} n^2 \tag{3.6.10}$$

with $n = 1, 2, , 3....$

These are the only values of the energy which allows solutions of the Schrödinger Equation 3.6.3 consistent with the boundary conditions in Equation 3.6.2. The integer n, called a **quantum number**, is appended as a subscript on E to label the allowed energy levels. Negative values of n add nothing new because the energies in Equation 3.6.10 depends on n^2.

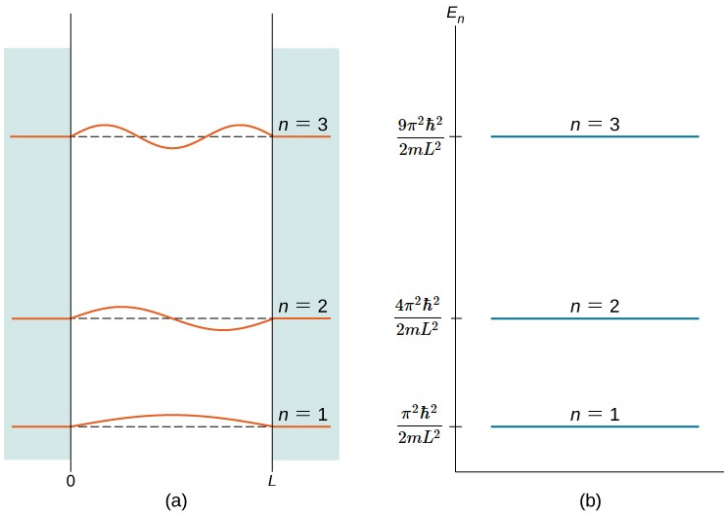


Figure 3.5.2 : A plot of $\psi_n(x)$ for the first four wavefunctions. (CC-BY 4.0; OpenStax).

Figure 3.5.2 shows part of the energy-level diagram for the particle in a box. The occurrence of discrete or quantized energy levels is characteristic of a bound system, that is, one confined to a finite region in space. For the free particle, the absence of confinement allowed an energy continuum. Note that, in both cases, the number of energy levels is infinite-denumerably infinite for the particle in a box, but nondenumerably infinite for the free particle.

The particle in a box assumes its lowest possible energy when $n = 1$, namely

$$E_1 = \frac{h^2}{8mL^2} \tag{3.6.11}$$

The state of lowest energy for a quantum system is termed its *ground state*.

Zero Point Energy

An interesting point is that $E_1 > 0$, whereas the corresponding classical system would have a minimum energy of zero. This is a recurrent phenomenon in quantum mechanics. The residual energy of the ground state, that is, the energy in excess of the classical minimum, is known as **zero point energy**. In effect, the kinetic energy, hence the momentum, of a bound particle cannot be reduced to zero. The minimum value of momentum is found by equating E_1 to $p^2/2m$, giving $p_{min} = \pm h/2L$. This can be expressed as an *uncertainty* in momentum given by $\Delta p \approx h/L$. Coupling this with the uncertainty in position, $\Delta x \approx L$, from the size of the box, we can write

$$\Delta x \Delta p \approx h \tag{3.6.12}$$

This is in accord with the *Heisenberg uncertainty principle*.

The particle-in-a-box eigenfunctions are given by Equation 3.6.13, with $B = 0$ and $k = n\pi/L = a$, in accordance with Equation 3.6.9

$$\psi_n(x) = A \sin \frac{n\pi x}{L} \tag{3.6.13}$$

with $n = 1, 2, 3....$

These, like the energies, can be labeled by the quantum number n. The constant A, thus far arbitrary, can be adjusted so that $\psi_n(x)$ is normalized. The normalization condition is, in this case,

$$\int_0^a [\psi_n(x)]^2 \, dx = 1 \tag{3.6.14}$$

the integration running over the domain of the particle $0 \le x \le L$. Substituting Equation 3.6.13 into Equation 3.6.14,

$$A^2 \int_0^L \sin^2 \frac{n\pi x}{L} dx = A^2 \frac{L}{n\pi} \int_0^{n\pi} \sin^2 \theta \, d\theta \tag{3.6.15}$$

$$= A^2 \frac{L}{2} = 1 \tag{3.6.16}$$

We have made the substitution $\theta = n\pi x/L$ and used the fact that the average value of $\sin^2 \theta$ over an integral number of half wavelengths equals 1/2 (alternatively, one could refer to standard integral tables). From Equation 3.6.16, we can identify the general normalization constant

$$A = \sqrt{\frac{2}{L}}$$

for all values of n. Finally we can write the normalized eigenfunctions:

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L} \tag{3.6.17}$$

with $n = 1, 2, 3....$

The first few eigenfunctions and the corresponding probability distributions are plotted in Figure 3.5.3 . There is a close analogy between the states of this quantum system and the modes of vibration of a violin string. The patterns of standing waves on the string are, in fact, identical in form with the wavefunctions in Equation 3.6.17.

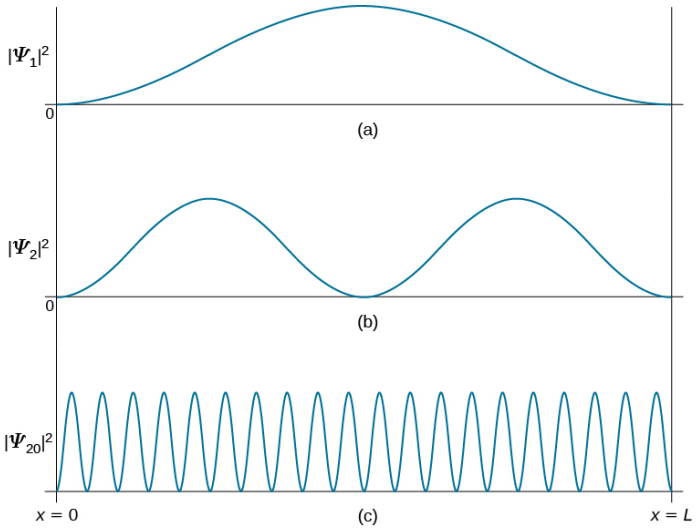


Figure 3.5.3 : The probability density distribution $|\psi_n(x)|^2$ for a quantum particle in a box for: (a) the ground state, $n = 1$; (b) the first excited state, $n = 2$; and, (c) the nineteenth excited state, $n = 20$. (CC-BY 4.0; OpenStax).

Nodes and Curvature

A significant feature of the particle-in-a-box quantum states is the occurrence of *nodes*. These are points, other than the two end points (which are fixed by the boundary conditions), at which the wavefunction vanishes. At a node there is exactly zero probability of finding the particle. The *n*th quantum state has, in fact, $n - 1$ nodes. It is generally true that the number of nodes increases with the energy of the quantum state, which can be rationalized by the following qualitative argument. As the number of nodes increases, so does the number and the steepness of the 'wiggles' in the wavefunction. It's like skiing down a slalom course. Accordingly, the average curvature, given by the second derivative, must increase. But the second derivative is proportional to the kinetic energy operator. Therefore, the more nodes, the higher the energy. This will prove to be an invaluable guide in more complex quantum systems.

Example 3.5.1 : Excited State Probabilities

For a particle in a one-dimensional box of length L, the second excited state wavefunction ($n = 3$) is

$$\psi_3 = \sqrt{\frac{2}{L}} \sin \frac{3\pi x}{L}$$

a. What is the probability that the particle is in the left half of the box?
b. What is the probability that the particle is in the middle third of the box?

Strategy

Probability that the particle will be found between a and b is

$$P(a,b) = \int_a^b \psi^2 \, dx$$

For this problem,

$$\psi_3 = \sqrt{\frac{2}{L}} \sin \frac{3\pi x}{L}$$

therefore,

$$P(a,b) = \frac{2}{L}\int_a^b \sin^2\frac{3\pi x}{L}\,dx$$
$$= \frac{2}{L}\left(\frac{x}{2} - \frac{L\sin\left(\frac{6\pi x}{L}\right)}{12\pi}\right)\Bigg|_a^b$$
$$= \frac{b-a}{L} - \frac{1}{6\pi}\left[\sin\left(\frac{6\pi b}{L}\right) - \sin\left(\frac{6\pi a}{L}\right)\right]$$

Solution a

The probability that the particle is in the left half of the box is

$$P\left(0,\frac{L}{2}\right) = \frac{\frac{L}{2}-0}{L} - \frac{1}{6\pi}\left[\sin\left(\frac{6\pi\left(\frac{L}{2}\right)}{L}\right) - \sin\left(\frac{6\pi(0)}{L}\right)\right]$$
$$= \frac{1}{2}$$

Solution b

The probability that the particle is in the middle third of the box

$$P\left(\frac{L}{3},\frac{2L}{3}\right) = \frac{\frac{2L}{3}-\frac{L}{3}}{L} - \frac{1}{6\pi}\left[\sin\frac{6\pi\left(\frac{2L}{3}\right)}{L} - \sin\frac{6\pi\left(\frac{L}{3}\right)}{L}\right]$$
$$= \frac{1}{3}$$

Exercise 3.5.1 : Ground State Probability

For a particle in a one-dimensional box, the ground state wavefunction is

$$\psi_1 = \sqrt{\frac{2}{L}}\sin\frac{x\pi}{L}$$

What is the probability that the particle is in the left half of the box in the ground state?

Time Dependence and Complex Nature of Wavefunctions

Recall that the time-dependence of the wavefunction with time-independent potential was discussed in Section 3.1 and is expressed as

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

so for the particle in a box, these are

$$\psi_n(x) = \sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}e^{-iE_nt/\hbar} \tag{3.6.18}$$

with E_n given by Equation 3.6.10.

The phase part of Equation 3.6.18 can be expanded into a real part and a complex components. So the total wavefunction for a particle in a box is

$$\Psi(x,t) = \underbrace{\left(\sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}\right)\left(\cos\frac{E_nt}{\hbar}\right)}_{\text{real part}} - \underbrace{i\left(\sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}\right)\left(\sin\frac{E_nt}{\hbar}\right)}_{\text{imaginary part}}$$

which can be simplified (slightly) to

$$\Psi(x,t) = \underbrace{\left(\sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}\right)\left(\cos\frac{E_n t}{\hbar}\right)}_{\text{real part}} - i\underbrace{\left(\sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}\right)\left(\cos\frac{E_n t}{\hbar} - \frac{\pi}{2}\right)}_{\text{imaginary part}}$$

As discussed previously, the imaginary part of the total wavefunction oscillates out of phase by $\pi/2$ with respect to the real part (we call this "out of phase"). This is demonstrated in the time-dependent behavior of the first three eigenfunctions in Figure 3.5.4 .

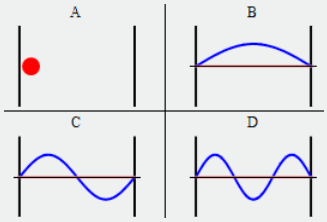


Figure 3.5.4 : Some trajectories of a particle in a box according to Newton's laws of classical mechanics (A), and according to the Schrödinger equation of quantum mechanics (B-D). In (B-D), the horizontal axis is position, and the vertical axis is the real part (blue) and imaginary part (red) of the wavefunction. The states (B, C, D) are energy eigenstates. (Public Domain; Sbyrnes321 via Wikipedia).

Note that as n increased, the energy of the wavefunction increases (Equation 3.6.10) and both the number of nodes and antinodes increase and the frequency of oscillation of the wavefunction increases.

> *It is generally true in quantum systems (not just for particles in boxes) that the number of nodes in a wavefunction increases with the energy of the quantum state.*

Orthonormality of the Eigenstates

Another important property of the eigenfunctions in Equation 3.6.17 applies to the integral over a product of two *different* eigenfunctions (Equation 3.6.17). It is easy to see from Figure 3.5.5 that the integral

$$\int_0^L \psi_2(x)\psi_1(x)dx = 0 \tag{3.6.19}$$

The integral in Equation 3.6.19 is zero when the two eignestates differ and when integrated over the entire range of the system (from $-\infty$ to ∞ for a 1-D particle in the box, but this can be a narrow region of 0 to L since the eigenstates are zero outside of the box).

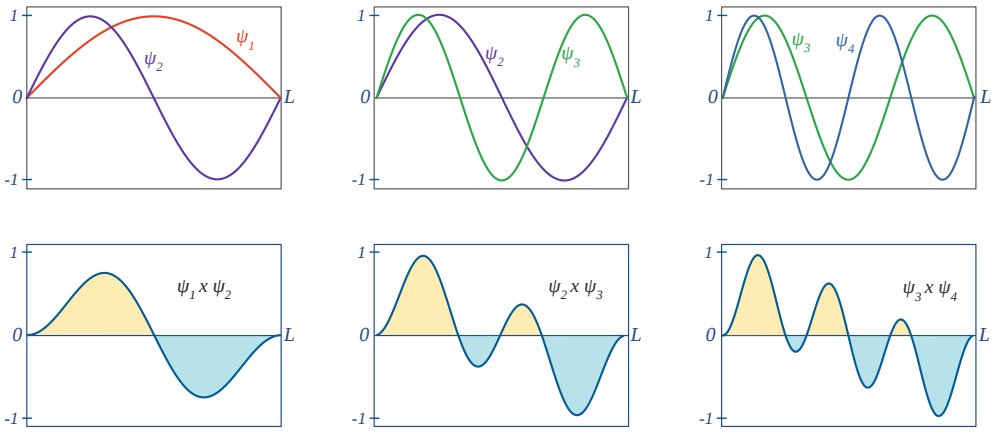


Figure 3.5.5 : (Top Row): Select plots of the $n = 1$ (red curves), $n = 2$ (purple curves), $n = 3$ (green curves), and $n = 4$ (blue curves) eigenfunctions. (Bottom Row): Product of different eigenstates with positive (tan) and negative (cyan) areas emphasized. Note that the negative and positive areas perfectly cancel when added together. Created via fooplot (CC BY-NC; Ümit Kaya via LibreTexts)

To prove Equation 3.6.19 for a particle in a box model, we can use the trigonometric identity

$$\sin\alpha\,\sin\beta = \frac{1}{2}\left[\cos(\alpha-\beta)-\cos(\alpha+\beta)\right] \tag{3.6.20}$$

to show that

$$\int_0^L \psi_m(x)\psi_n(x)dx = 0 \quad if \quad m \neq n \tag{3.6.21}$$

This property is called **orthogonality**. We will show in the next chapter, that this is a general result from quantum-mechanical eigenfunctions. The normalization (Equation 3.6.19) together with the orthogonality (Equation 3.6.21) can be combined into a single relationship

$$\int_0^L \psi_m(x)\psi_n(x)dx = \delta_{mn} \tag{3.6.22}$$

In terms of the *Kronecker delta*

$$\delta_{mn} = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases} \tag{3.6.23}$$

A set of functions $\{\psi_n\}$ which obeys Equation 3.6.22 is called *orthonormal*.

Example 3.5.2

Evaluate

a. $\langle\psi_3|\psi_3\rangle$
b. $\langle\psi_4|\psi_4\rangle$
c. $\langle\psi_3|\psi_4\rangle$
d. $\langle\psi_4|\psi_3\rangle$

for the normalized wavefunctions:

$$|\psi_3\rangle = \sqrt{\frac{2}{L}}\sin\frac{3\pi x}{L}$$

and

$$|\psi_4\rangle = \sqrt{\frac{2}{L}}\sin\frac{4\pi x}{L}$$

Strategy

These are four different integrals and we can solve them directly or use orthonormality (Equation 3.6.23) to evaluate.

Solution a

$$\begin{aligned}\langle\psi_3|\psi_3\rangle &= \int_{-\infty}^{+\infty}\left(\sqrt{\frac{2}{L}}\sin\frac{3\pi x}{L}\right)\left(\sqrt{\frac{2}{L}}\sin\frac{3\pi x}{L}\right)dx \\ &= \frac{2}{L}\int_{-\infty}^{+\infty}\sin^2\frac{3\pi x}{L}\,dx\end{aligned}$$

This is an integration over an even function, so it cannot be tossed out via symmetry. We can use the Trigonometry relationship in Equation 3.6.20 to get

$$\frac{2}{L}\int_{-\infty}^{+\infty}\sin^2\frac{3\pi x}{L}\,dx = \frac{2}{L}\int_{-\infty}^{+\infty}\frac{1}{2}\left(1-\cos\frac{6\pi x}{L}\right)dx$$

and we can continue the fun. However, there is no need. Since the we can recognize that $\langle\psi_3|\psi_3\rangle$ is 1 by the normalization criteria which is folded into the orthonormal criteria (Equation 3.6.23).

Therefore $\langle\psi_3|\psi_3\rangle = 1$.

Solution b

$$\langle \psi_4 | \psi_4 \rangle = \int_{-\infty}^{+\infty} \left(\sqrt{\frac{2}{L}} \sin \frac{4\pi x}{L} \right) \left(\sqrt{\frac{2}{L}} \sin \frac{4\pi x}{L} \right) dx$$
$$= \frac{2}{L} \int_{-\infty}^{+\infty} \sin^2 \frac{4\pi x}{L} \, dx$$

We can expand and solve, but again there is no need. The wavefunctions are normalized therefore $\langle \psi_4 | \psi_4 \rangle = 1$.

Solution c

$$\langle \psi_3 | \psi_4 \rangle = \int_{-\infty}^{+\infty} \left(\sqrt{\frac{2}{L}} \sin \frac{3\pi x}{L} \right) \left(\sqrt{\frac{2}{L}} \sin \frac{4\pi x}{L} \right) dx$$
$$= \frac{2}{L} \int_{-\infty}^{+\infty} \sin \left(\frac{3\pi x}{L} \right) \sin \left(\frac{4\pi x}{L} \right) dx$$

We can expand this integral and evaluate, but since the integrand is odd (and even function times an odd), this integral is zero. Alternatively, we can use the orthogonality criteria into the greater orthonormal criteria (Equation 3.6.23).

Solution d

$$\langle \psi_4 | \psi_3 \rangle = \int_{-\infty}^{+\infty} \left(\sqrt{\frac{2}{L}} \sin \frac{4\pi x}{L} \right) \left(\sqrt{\frac{2}{L}} \sin \frac{3\pi x}{L} \right) dx$$
$$= \frac{2}{L} \int_{-\infty}^{+\infty} \sin \left(\frac{4\pi x}{L} \right) \sin \left(\frac{3\pi x}{L} \right) dx$$

We can expand this integral and evaluate, but since the integrand is odd (and even function times an odd function), this integral is zero. Alternatively, we can use the orthogonality criteria into the greater orthonormal criteria (Equation 3.6.23).

However, since the wavefunctions are real, then

$$\langle \psi_4 | \psi_3 \rangle = \langle \psi_3 | \psi_4 \rangle$$

which also means

$$\langle \psi_4 | \psi_3 \rangle = 0$$

from the results of section c.

3.6: The Energy of a Particle in a Box is Quantized is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.5: The Energy of a Particle in a Box is Quantized** by S. M. Blinder is licensed CC BY-NC-SA 4.0.

3.7: Wavefunctions Must Be Normalized

Learning Objectives

- Calculate the probability of an event from the wavefunction
- Understand the utility and importance of normalizing wavefunctions
- Demonstrate how to normalize an arbitrary wavefunction

Extracting Probabilities

Since wavefunctions can in general be complex functions, the physical significance of wavefunctions cannot be found from the functions themselves because the $\sqrt{-1}$ is not a property of the physical world. Rather, the physical significance is found in the product of the wavefunction and its complex conjugate, i.e. the absolute square of the wavefunction, which also is called the **square of the modulus**.

$$\Psi^*(r,t)\Psi(r,t) = |\Psi(r,t)|^2 \tag{3.7.1}$$

where r is a vector specifying a point in three-dimensional space. The square is used, rather than the modulus itself, just like the intensity of a light wave depends on the square of the electric field. Remember that the Born interpretation is that $\psi^*(r_i)\psi(r_i)\,d\tau$ is the **probability** that the electron is in the volume $d\tau$ located at r_i. The Born interpretation therefore calls the wavefunction the probability amplitude, the absolute square of the wavefunction is called the **probability density**, and the probability density times a volume element in three-dimensional space ($d\tau$) is the probability.

Since the squared magnitude $|\psi|^2$ of the wavefunction of a particle can be interpreted as the probability density, then the probability for a one-dimensional wavefunction between the points $x = a$ and $x = b$ can be calculated by

$$P_{1D} = \int_a^b |\psi(x)|^2 \mathrm{d}x \tag{3.7.2}$$

This is just the area under the under the $|\psi|^2$ curve (Figure 3.6.1).

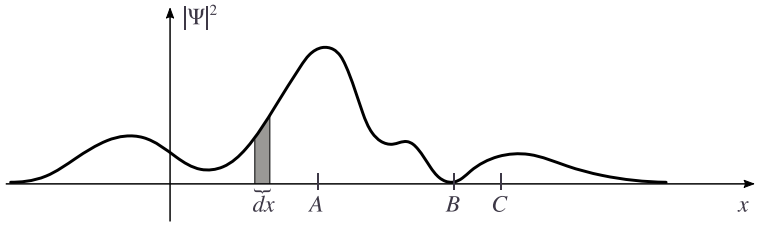


Figure 3.6.1 : The probability can be interpreted as an area under the probability density $|\psi|^2$. (CC BY-NC 4.0; Ümit Kaya via LibreTexts)

If the probability of a two-dimensional wavefunction is being evaluated, then Equation 3.7.2 will be amended to included a double integral:

$$P_{2D} = \iint_{a_1,a_2}^{b_1,b_2} |\psi(x,y)|^2 \mathrm{d}x\mathrm{d}y$$

and similarly a triple integral would be used for calculating probabilities of three-dimensional wavefunctions:

$$P_{3D} = \iiint_{a_1,a_2,a_3}^{b_1,b_2,b_3} |\psi(x,y,z)|^2 \mathrm{d}x\mathrm{d}y\mathrm{d}z$$

Example 3.6.1 : Probability of a Particle in a Box

Calculate the probability of finding an electron at $L/2$ in a box of infinite height within an interval ranging from $\frac{L}{2} - \frac{L}{200}$ to $\frac{L}{2} + \frac{L}{200}$ for the $n = 1$ and $n = 2$ states. Since the length of the interval, $L/100$, is small compared to L, you can get an approximate answer without explicitly integrating.

Solution

The wavefunction for the particle in a box is

$$\psi(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right)$$

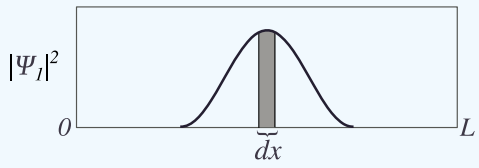
and the wavefunction for the $n = 1$ state is

$$\psi_{n=1} = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right)$$

From the interpretation that the wavefunction modulus squared is the probability density, we can establish the following integral to solve the problem (note the limits of integration)

$$|\psi_{n=1}|^2 = \frac{2}{L} \int_{\frac{99L}{200}}^{\frac{101L}{200}} \sin^2\left(\frac{\pi x}{L}\right) dx \tag{3.7.3}$$

We can solve this, but we can also recognize that Equation 3.7.3 is just calculating an area that can be approximated as the area of a rectangle with a height ($\frac{2}{L}\sin^2\left(\frac{\pi x}{L}\right)$ at $x = L/2$) and width $\Delta x = L/100$ (Figure 3.6.2).



(a)

Figure 3.6.2 : The probability can be interpreted as an area under the probability density $|\psi_1|^2$. (CC BY-NC; Ümit Kaya via LibreTexts)

This area can be computed:

$$\begin{aligned} |\psi_{n=1}|^2 &\approx \frac{2}{L} \sin^2\left(\frac{\pi(L/2)}{L}\right)^{1} \Delta x \\ &\approx \left(\frac{2}{L}\right)(L/100) \\ &\approx 1/50 = 0.02 \end{aligned}$$

Given that the wavefunction is sinusoidal, the actual probability of finding an electron within the given interval at $\frac{L}{2}$ should be slightly less because of the behavior of the sinusoid at $\frac{L}{2}$ is at its peak of the wavefunction (Figure 3.6.2).

The wavefunction for the $n = 2$ state

$$\psi_{n=2} = \sqrt{\frac{2}{L}} \sin\left(\frac{2\pi x}{L}\right)$$

so the integral that we need to construct and solve is

$$|\psi_{n=2}|^2 = \frac{2}{L} \int_{\frac{99L}{200}}^{\frac{101L}{200}} \sin^2\left(\frac{2\pi x}{L}\right) dx$$

We can use the same graphical interpretation as above, but using the probability density of the ψ_2 wavefunction (Figure 3.6.3).

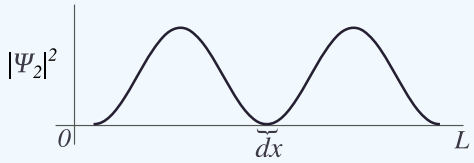


Figure 3.6.3 : The probability can be interpreted as an area under the probability density $|\psi_2|^2$. (CC BY-NC Copyright; Ümit Kaya via LibreTexts)

$$|\psi_{n=1}|^2 \approx \frac{2}{L}\sin^2\left(\frac{2\pi(L/2)}{L}\right)^{0}\Delta x$$
$$\approx 0$$

The probability of finding an electron in a box at $\frac{L}{2}$ for $n = 2$ is approximately zero.

? Exercise 3.6.1

Show that the square of the modulus of $\Psi(r,t) = \psi(r)e^{-i\omega t}$ is time independent. What insight regarding stationary states do you gain from this proof?

Solution

The square of the modulus of a wavefunction is $\Psi(r,t)^*\Psi(r,t)$ so for wavefunctions of this form, the square of the modulus is

$$\Psi(r,t)^*\Psi(r,t) = \psi(r)\,e^{+i\omega t}\,\psi(r)\,e^{-i\omega t}$$
$$= \psi(r)^2$$

Hence, there is no time dependence to the modulus of wavefunctions of this work, which from the probability interpretation of the wavefunction means that the probability density is time-independent.

Normalizing of the Wavefunction

A probability is a real number between 0 and 1, inclusive. An outcome of a measurement which has a probability 0 is an impossible outcome, whereas an outcome which has a probability 1 is a certain outcome. According to Equation 3.7.1, the probability of a measurement of x yielding a result between $-\infty$ and $+\infty$ is

$$P_{x\in-\infty:\infty}(t) = \int_{-\infty}^{\infty}|\psi(x,t)|^2dx. \tag{3.7.4}$$

However, a measurement of x *must* yield a value between $-\infty$ and $+\infty$, since the particle has to be located somewhere. It follows that $P_{x\in-\infty:\infty}(t) = 1$, or

$$\int_{-\infty}^{\infty}|\psi(x,t)|^2dx = 1, \tag{3.7.5}$$

which is generally known as the **normalization condition** for the wavefunction.

✓ Example 3.6.2 : Normalizing a Gaussian Wavepacket

Normalize the wavefunction of a Gaussian wave packet, centered on $x = x_o$ with characteristic width σ:

$$\psi(x) = \psi_0 e^{-(x-x_0)^2/(4\sigma^2)}. \tag{3.7.6}$$

Solution

To determine the normalization constant ψ_0, we simply substitute Equation 3.7.6 into Equation 3.7.5, to obtain

$$|\psi_0|^2\int_{-\infty}^{\infty}e^{-(x-x_0)^2/(2\sigma^2)}dx = 1.$$

Changing the variable of integration to $y = (x - x_0)/(\sqrt{2}\sigma)$, we get

$$|\psi_0|^2 \sqrt{2}\sigma \int_{-\infty}^{\infty} \mathrm{e}^{-y^2} dy = 1.$$

However, from an integral table we know

$$\int_{-\infty}^{\infty} \mathrm{e}^{-y^2} dy = \sqrt{\pi},$$

which implies that

$$|\psi_0|^2 = \frac{1}{(2\pi\sigma^2)^{1/2}}.$$

Hence, a general normalized Gaussian wavefunction takes the form

$$\psi(x) = \frac{e^{i\phi}}{(2\pi\sigma^2)^{1/4}} e^{-(x-x_0)^2/(4\sigma^2)}$$

where ϕ is an arbitrary real phase-angle.

? Exercise 3.6.2

Normalize this wavefunction for a particle in a harmonic well:

$$\psi = xe^{-x^2}$$

Time Dependence to the Wavefunction

Now, it is important to demonstrate that if a wavefunction is initially normalized then it stays normalized as it evolves in time according to the time-dependent Schrödinger's equation. If this is not the case then the probability interpretation of the wavefunction is untenable, since it does not make sense for the probability that a measurement of x yields *any* possible outcome (which is, manifestly, unity) to change in time. Hence, we require that

$$\frac{d}{dt} \int_{-\infty}^{\infty} |\psi(x,t)|^2 dx = 0$$

for wavefunctions satisfying the time-dependent Schrödinger's equation (this results from the time-dependent Schrödinger's equation and Equation 3.7.5). The above equation gives

$$\frac{d}{dt} \int_{-\infty}^{\infty} \psi^* \psi \, dx = \int_{-\infty}^{\infty} \left(\frac{\partial \psi^*}{\partial t} \psi + \psi^* \frac{\partial \psi}{\partial t} \right) dx = 0. \tag{3.7.7}$$

Now, multiplying Schrödinger's equation by $\psi^*/(\mathrm{i}\hbar)$, we obtain

$$\psi^* \frac{\partial \psi}{\partial t} = \frac{\mathrm{i}}{2m} \psi^* \frac{\partial^2 \psi}{\partial x^2} - \frac{\mathrm{i}}{\hbar} V |\psi|^2 \tag{3.7.8}$$

The complex conjugate of this expression yields

$$\psi \frac{\partial \psi^*}{\partial t} = -\frac{\mathrm{i}}{2m} \psi \frac{\partial^2 \psi^*}{\partial x^2} + \frac{\mathrm{i}}{\hbar} V |\psi|^2 \tag{3.7.9}$$

since

- $(AB)^* = A^* B^*$,
- $A^{**} = A$, and
- $i^* = -i$.

Summing Equation 3.7.8 and 3.7.9 results in

$$\frac{\partial \psi^*}{\partial t}\psi + \psi^*\frac{\partial \psi}{\partial t} = \frac{\mathrm{i}}{2m}\left(\psi^*\frac{\partial^2 \psi}{\partial x^2} - \psi\frac{\partial^2 \psi^*}{\partial x^2}\right) \tag{3.7.10}$$

$$= \frac{\mathrm{i}}{2m}\frac{\partial}{\partial x}\left(\psi^*\frac{\partial \psi}{\partial x} - \psi\frac{\partial \psi^*}{\partial x}\right). \tag{3.7.11}$$

Equations 3.7.7 and 3.7.11 can be combined to produce

$$\frac{d}{dt}\int_{-\infty}^{\infty}|\psi|^2 dx = \frac{\mathrm{i}}{2m}\left[\psi^*\frac{\partial \psi}{\partial x} - \psi\frac{\partial \psi^*}{\partial x}\right]_{-\infty}^{\infty} = 0. \tag{3.7.12}$$

The above equation is satisfied provided the wavefunction converges

$$\lim_{|x|\to\infty}|\psi| = 0 \tag{3.7.13}$$

However, this is a necessary condition for the integral on the left-hand side of Equation 3.7.5 to converge. Hence, we conclude that all wavefunctions which are *square-integrable* [i.e., are such that the integral in Equation 3.7.5 converges] have the property that if the normalization condition Equation 3.7.5 is satisfied at one instant in time then it is satisfied at all subsequent times.

Not all Wavefunctions can be Normalized

Not all wavefunctions can be normalized according to the scheme set out in Equation 3.7.5. For instance, a planewave wavefunction for a quantum free particle

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

is not square-integrable, and, thus, cannot be normalized. For such wavefunctions, the best we can say is that

$$P_{x\in a:b}(t) \propto \int_a^b |\Psi(x,t)|^2 dx.$$

In the following, all wavefunctions are assumed to be square-integrable and normalized, unless otherwise stated.

3.7: Wavefunctions Must Be Normalized is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.6: Wavefunctions Must Be Normalized** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

3.8: The Average Momentum of a Particle in a Box is Zero

Learning Objectives

- Calculate the expectation value for a measurement
- Apply the expectation value concept to calculate average properties of a participle in a box model
- Understand the origin of a zero-point energy/zero-point motion.
- Extend the concept of orthogonality from vectors to mathematical functions (and wavefunctions).

Now that we have mathematical expressions for the wavefunctions and energies for the particle-in-a-box, we can answer a number of interesting questions. The answers to these questions use quantum mechanics to predict some important and general properties for electrons, atoms, molecules, gases, liquids, and solids. Key to addressing these questions is the formulation and use of expectation values. This is demonstrated below and used in the context of evaluating average properties (momentum of the particle in a box for the case below).

Classical Expectation Values

The expectation value is the probabilistic expected value of the result (measurement) of an experiment. It is not the most probable value of a measurement; indeed the expectation value may even have zero probability of occurring. The expected value (or expectation, mathematical expectation, mean, or first moment) refers to the value of a variable one would "expect" to find if one could repeat the random variable process an infinite number of times and take the average of the values obtained. More formally, the expected value is a weighted average of all possible values.

Example 3.7.1 : Classical Expectation Value of Exam Scores (a discretized example)

A classical example is calculating the expectation value (i.e. average) of the exam grades in the class. For example if the class scores for an exam were

65	67	94	43	67	76	94	67

The discrete way is to sum up all scores and divide by the number of students:

$$\langle s \rangle = \frac{\sum_i^N s(i)}{N} \tag{3.8.1}$$

which of this example of scores is

$$\begin{aligned} \langle s \rangle &= \frac{65+67+94+43+67+76+94+76}{8} \\ &= 71.625 \end{aligned}$$

Notice that the average is not an allowable score on an individual exam. Equation 3.8.1 can be rewritten with "probability" or "probability weights"

$$\langle s \rangle = \sum_i^N s(i) P_s(i) \tag{3.8.2}$$

where $P_s(i)$ is the probability of observing a score of s. This is just the number of times it occurs in a dataset divided by the number of elements in that data set. Applying Equation 3.8.2 to the set of scores, we need to calculate these weights:

Score	65	67	94	43	76
P_s	1/8	3/8	2/8	1/8	1/8

As with all probabilities, the sum of all probabilities possible must be one. These confirm that for the weights here:

$$\frac{1}{8}+\frac{3}{8}+\frac{2}{8}+\frac{1}{8}+\frac{1}{8}=\frac{8}{8}=1$$

This is the discretized "normalization" criterion (the same as why we normalize wavefunctions).

So, now we can use Equation 3.8.2 properly

$$\begin{aligned}\langle s\rangle &= 65\times\frac{1}{8}+67\times\frac{3}{8}+94\times\frac{2}{8}+43\times\frac{1}{8}+76\times\frac{1}{8}\\ &= 71.625\end{aligned}$$

Hence, Equation 3.8.2 gives the same result, as expected, from Equation 3.8.1.

Quantum Expectation Values

The extension of the classical expectation (average) approach in Example 3.7.1 using Equation 3.8.2 to evaluating quantum mechanical expectation values requires three small changes:

1. Switch from descretized to continuous variables
2. Substitute the wavefunction squared for the probability weights (i.e., the probability distribution)
3. Use an operator instead of the scalar

Hence, the quantum mechanical expectation value $\langle o\rangle$ for an observable, o, associated with an operator, $\hat{O}$, is given by

$$\langle o\rangle = \int_{-\infty}^{+\infty}\psi^*\hat{O}\psi\,dx \tag{3.8.3}$$

where x is the range of space that is integrated over (i.e., an integration over all possible probabilities). The expectation value changes as the wavefunction changes and the operator used (i.e, which observable you are averaging over).

> *In general, changing the wavefunction changes the expectation value for that operator for a state defined by that wavefunction.*

Average Energy of a Particle in a Box

If we generalize this conclusion, such integrals give the average value for any physical quantity by using the operator corresponding to that physical observable in the integral in Equation 3.8.3. In the equation below, the symbol $\langle H\rangle$ is used to denote the average value for the total energy.

$$\langle H\rangle = \int_{-\infty}^{\infty}\psi^*(x)\hat{H}\psi(x)dx \tag{3.8.4}$$

$$= \int_{-\infty}^{\infty}\psi^*(x)\hat{KE}\psi(x)dx+\int_{-\infty}^{\infty}\psi^*(x)\hat{V}\psi(x)dx \tag{3.8.5}$$

$$= \underbrace{\int_{-\infty}^{\infty}\psi^*(x)\left(\frac{-\hbar^2}{2m}\right)\frac{\partial^2}{\partial x^2}\psi(x)dx}_{\text{average kinetic energy}}+\underbrace{\int_{-\infty}^{\infty}\psi^*(x)\hat{V}(x)\psi(x)dx}_{\text{average potential energy}} \tag{3.8.6}$$

The Hamiltonian operator consists of a kinetic energy term and a potential energy term. The kinetic energy operator involves differentiation of the wavefunction to the right of it. This step must be completed before multiplying by the complex conjugate of the wavefunction. The potential energy, however, usually depends only on position and not momentum (i.e., it involves conservative forces). The potential energy operator therefore only involves the coordinates of a particle and does not involve differentiation. For this reason we do not need to use a caret over V in Equation 3.8.6.

Equation 3.8.6 can be simplified

$$\langle H\rangle = \langle KE\rangle+\langle V\rangle \tag{3.8.7}$$

The potential energy integral then involves only products of functions, and the order of multiplication does not affect the result, e.g. 6×4 = 4×6 = 24. This property is called the **commutative property**. The average potential energy therefore can be written as

$$\langle V \rangle = \int_{-\infty}^{\infty} V(x)\psi^*(x)\psi(x)dx \tag{3.8.8}$$

This integral is telling us to take the probability that the particle is in the interval dx at x, which is $\psi^*(x)\psi(x)dx$, multiply this probability by the potential energy at x, and sum (i.e., integrate) over all possible values of x. This procedure is just the way to calculate the average potential energy $\langle V \rangle$ of the particle.

Exercise 3.7.2 : Particle in Box

Evaluate the two integrals in Equation 3.8.6 for the PIB wavefunction $\psi(x) = \sqrt{\frac{2}{L}}\sin(kx)$ with the potential function $V(x) = 0$ from 0 to the length of a box L with $k = \frac{\pi}{L}$.

Solution

The *average kinetic energy* is

$$\begin{aligned}
\langle KE \rangle &= \int_0^L \left(\sqrt{\frac{2}{L}}\right)\sin(kx)\left(\frac{-\hbar^2}{2m}\right)\frac{\partial^2}{\partial x^2}\left(\sqrt{\frac{2}{L}}\right)\sin(kx)dx \\
&= \left(\frac{2}{L}\right)\int_0^L \sin(kx)\left(\frac{-\hbar^2}{2m}\right)\frac{\partial}{\partial x}\cos(kx)(k)dx \\
&= \left(\frac{2}{L}\right)\int_0^L \sin(kx)\left(\frac{-\hbar^2}{2m}\right)\sin(kx)(k)(-k)dx \\
&= \left(\frac{2}{L}\right)\left(\frac{k^2\hbar^2}{2m}\right)\int_0^L \sin^2(kx)dx
\end{aligned}$$

We can solve this intergral using the standard half-angle representation from an integral table. Or we can recognize that we already did this integral when we normalized the PIB wavefunction by rewriting this integral:

$$\begin{aligned}
\langle KE \rangle &= \left(\frac{2}{L}\right)\left(\frac{k^2\hbar^2}{2m}\right)\int_0^L \sin^2(kx)dx \\
&= \left(\frac{k^2\hbar^2}{2m}\right)\int_0^L \left(\frac{2}{L}\right)\sin^2(kx)dx \\
&= \left(\frac{k^2\hbar^2}{2m}\right)\int_0^L \psi^*(x)\psi(x)dx \quad (\to 1) \\
&= \frac{k^2\hbar^2}{2m}
\end{aligned}$$

Thus, the average value for the total energy of this particular system is

$$\langle KE \rangle = \frac{k^2\hbar^2}{2m} = \frac{\pi^2\hbar^2}{2mL^2}$$

Hence, the average kinetic energy of the wavefunction is dependent on the n quantum number

The *average potential energy* is

$$\langle V \rangle = \int_{-\infty}^{\infty} \sin(kx) 0 \sin(kx) dx = 0$$

Thus, the average potential energy of the PIB is 0 irrespective of the wavefunction.

Hence via Equation 3.8.7 for this system and set of wavefunctions

$$\langle H \rangle = \frac{\pi^2 \hbar^2}{2mL^2}$$

This is the same result obtained from solving the eigenvalue equation for the PIB. However, if the wavefunctions used were NOT eigenstates of energy, then we cannot use the eigenvalue approach and need to rely on the expectation values to describe the energy of the system.

What is the lowest energy for a particle in a box? The lowest energy level is E_1, and it is important to recognize that this lowest energy of a particle in a box is not zero. This finite energy is called the **zero-point energy**, and the motion associated with this energy is called the zero-point motion. Any system that is restricted to some region of space is said to be bound. The zero-point energy and motion are manifestations of the wave properties and the Heisenberg Uncertainty Principle, and are general properties of bound quantum mechanical systems.

? Exercise 3.7.2 : Progressing to the Classical Limit

What happens to the energy level spacing for a particle-in-a-box when mL^2 becomes much larger than $\hbar^2$? What does this result imply about the relevance of quantization of energy to baseballs in a box between the pitching mound and home plate? What implications does quantum mechanics have for the game of baseball in a world where $\hbar$ is so large that baseballs exhibit quantum effects?

The first derivative of a function is the rate of change of the function, and the second derivative is the rate of change in the rate of change, also known as the curvature. A function with a large second derivative is changing very rapidly. Since the second derivative of the wavefunction occurs in the Hamiltonian operator that is used to calculate the energy by using the Schrödinger equation, a wavefunction that has sharper curvatures than another, i.e. larger second derivatives, should correspond to a state having a higher energy. A wavefunction with more nodes than another over the same region of space must have sharper curvatures and larger second derivatives, and therefore should correspond to a higher energy state.

? Exercise 3.7.3 : Nodes and Energies

Identify a relationship between the number of nodes in a wavefunction and its energy by examining the graphs you made above. A node is the point where the amplitude passes through zero. What does the presence of many nodes mean about the shape of the wavefunction?

Average Position of a Particle in a Box

We can calculate the most probable position of the particle from knowledge of probability distribution, $\psi^* \psi$. For the ground-state particle in a box wavefunction with $n = 1$ (Figure 3.8.1*a*)

$$\psi_{n=1} = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right) \tag{3.8.9}$$

This state has the following probability distribution (Figure 3.8.1*b*):

$$\psi^*_{n=1} \psi_{n=1} = \frac{2}{L} \sin^2\left(\frac{\pi x}{L}\right)$$

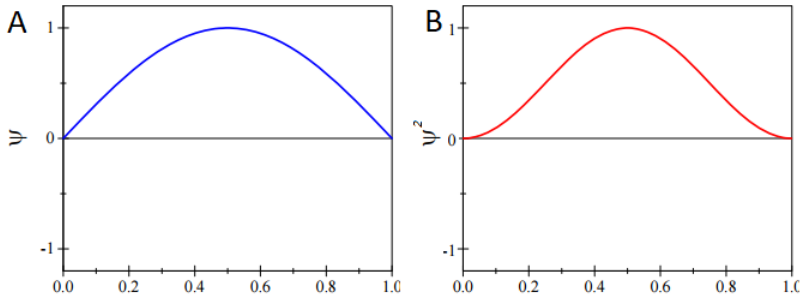


Figure 3.7.1 : (left) The ground-state ($n = 1$) wavefunction for a particle in a box. (right) The ground-state ($n = 1$) probability for a particle in a box.

The expectation value for position with the $\hat{x} = x$ operation for any wavefunction (Equation 3.8.3) is

$$\langle x \rangle = \int_{-\infty}^{+\infty} \psi^* x \psi \, dx$$

which for the ground-state wavefunction (Equation 3.8.9) shown in Figure 3.7.1 is

$$\langle x \rangle = \int_{-\infty}^{+\infty} \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right) x \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right) dx \tag{3.8.10}$$

$$= \frac{2}{L} \int_{-\infty}^{+\infty} x \sin^2\left(\frac{\pi x}{L}\right) \tag{3.8.11}$$

Solution by Inspection

Without even having to evaluate Equation 3.8.11, we can get the expectation value from simply inspecting $\psi^*_{n=1}\psi_{n=1}$ in (Figure 3.8.1;). This is a symmetric distribution around the center of the box ($L/2$) so it is just as likely to be found in the left half than the right half. Moreover, specifically at any point a fixed distance away from the mean, i.e.

$$\psi^*_{n=1}\psi_{n=1}(L/2 + \Delta x) = \psi^*_{n=1}\psi_{n=1}(L/2 - \Delta x)$$

Therefore, the particle is most likely to be found at the center of the box. So

$$\langle x \rangle = \frac{L}{2}$$

Exercise 3.7.4

Use the general form of the particle-in-a-box wavefunction for any n to find the mathematical expression for the position expectation value $\langle x \rangle$ for a box of length L. How does $\langle x \rangle$ depend on n?

Average Momentum of a Particle in a Box

What is the average momentum of a particle in the box? We start with Equation 3.8.3 and use the momentum operator

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x} \tag{3.8.12}$$

We note that the particle-in-a-box wavefunctions are **not** eigenfunctions of the momentum operator (Exercise 3.7.4). However, this does not mean that Equation 3.8.3 is inapplicable as Example 3.7.2 demonstrates.

Example 3.7.3 : The Average Momentum of a Particle in a Box is Zero

Even though the wavefunctions are not momentum eigenfunctions, we can calculate the expectation value for the momentum. Show that the expectation or average value for the momentum of an electron in the box is zero in every state (i.e., arbitrary values of n).

Strategy

First write the expectation value integral (Equation 3.8.3) with the momentum operator. Then insert the expression for the wavefunction and evaluate the integral as shown here.

Solution

$$\begin{aligned}\langle p\rangle &= \int_0^L \psi_n^*(x)\left(-i\hbar\frac{d}{dx}\right)\psi_n(x)dx \\ &= \int_0^L \left(\frac{2}{L}\right)^{1/2}\sin\left(\frac{n\pi x}{L}\right)\left(-i\hbar\frac{d}{dx}\right)\left(\frac{2}{L}\right)^{1/2}\sin\left(\frac{n\pi x}{L}\right)dx \\ &= -i\hbar\left(\frac{2}{L}\right)\int_0^L \sin\left(\frac{n\pi x}{L}\right)\left(\frac{d}{dx}\right)\sin\left(\frac{n\pi x}{L}\right)dx \\ &= -i\hbar\left(\frac{2}{L}\right)\left(\frac{n\pi}{L}\right)\int_0^L \sin\left(\frac{n\pi x}{L}\right)\cos\left(\frac{n\pi x}{L}\right)dx \\ &= 0\end{aligned}$$

Note that this makes sense since the particles spends an equal amount of time traveling in the $+x$ and $-x$ direction.

Interpretation

It may seem that this means the particle in a box does not have any momentum, which is incorrect because we know the energy is never zero. In fact, the energy that we obtained for the particle-in-a-box is entirely kinetic energy because we set the potential energy at 0. Since the kinetic energy is the momentum squared divided by twice the mass, it is easy to understand how the average momentum can be zero and the kinetic energy finite

? Exercise 3.7.5

Show that the particle-in-a-box wavefunctions are not eigenfunctions of the momentum operator (Equation 3.8.12).

? Exercise 3.7.6

Write the particle-in-a-box wavefunction as a normalized linear combination of the momentum eigenfunctions e^{ikx} and e^{-ikx} by using Euler's formula. Show that the momentum eigenvalues for these two functions are $p = +\hbar k$ and $-\hbar k$.

The interpretation of the results of Exercise 3.7.6 is physically interesting. The exponential wavefunctions in the linear combination for the sine function represent the two opposite directions in which the electron can move. One exponential term represents movement to the left and the other term represents movement to the right (Figure 3.7.2).

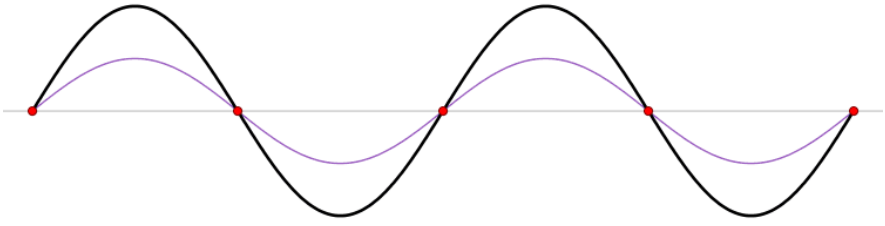


Figure 3.7.2 : Two waves traveling in opposite directions across the same medium combine linearly. In this animation, both waves have the same wavelength and the sum of amplitudes results in a standing wave. (Public Domain; LucasVB).

> *The electrons are moving, they have kinetic energy and momentum, yet the average momentum is zero.*

Did we just Violate the Uncertainty Principle?

Does the fact that the average momentum of an electron is zero and the average position is $L/2$ violate the Heisenberg Uncertainty Principle? No, because the Heisenberg Uncertainty Principle pertains to the uncertainty in the momentum and in the position, not to the average values. Quantitative values for these uncertainties can be obtained to compare with the limit set by the Heisenberg Uncertainty Principle for the product of the uncertainties in the momentum and position. However, to do this we need a quantitative definition of uncertainty, which is discussed in the following Section.

Orthogonality

In vector calculus, *orthogonality* is the relation of two lines at right angles to one another (i.e., perpendicularity), but is generalized into n dimensions via zero amplitude "dot products" or "inner products." Hence, orthogonality is thought of as describing **non-overlapping**, **uncorrelated**, or **independent** objects of some kind. The concept of orthogonality extends to functions (wavefunctions or otherwise) too. Two functions ψ_A and ψ_B are said to be orthogonal if

$$\int_{allspace} \psi_A^* \psi_B d\tau = 0 \tag{3.8.13}$$

In general, eigenfunctions of a quantum mechanical operator with different eigenvalues are orthogonal. Are the eigenfunctions of the particle-in-a-box Hamiltonian orthogonal?

Exercise 3.7.7

Evaluate the integral $\int \psi_1^* \psi_3 dx$ for all possible pairs of particle-in-a-box eigenfunctions from $n = 1$ to $n = 3$ (use symmetry arguments whenever possible) and explain what the results say about orthogonality of the functions.

3.8: The Average Momentum of a Particle in a Box is Zero is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.7: The Average Momentum of a Particle in a Box is Zero** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

3.9B: Particle in a Finite Box and Tunneling (optional)

The finite potential well is an extension of the infinite potential well from the previous section. The main difference between these two systems is that now the particle has a non-zero probability of finding itself outside the well, although its kinetic energy is less than that required, according to classical mechanics, for scaling the potential barrier [4]. This type of problem is more realistic, but more difficult to solve due to the yielding of transcendental equations. The particle is again confined to a box, but one which has finite, not infinite, potential walls. We consider a potential well of height V0.

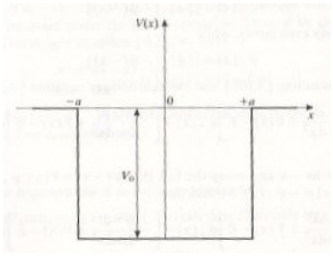


Figure 2.1 *Square well with finite potential. Here, a=L.*

We have the following potential, V(x), given by the boundary conditions shown in Figure 2.1:

$$V(x) = \begin{cases} -V_0 & |x| < L \\ 0 & |x| > L \end{cases}$$

In this example, the origin of the x-axis was chosen at the center of the well. By doing so, the potential is symmetric about x=0, giving rise to parity (Note: this could also be applied to a symmetric infinite wells). For the finite well, two cases must be distinguished, corresponding to positive or negative values of the energy E [1]. It is possible for the particle to be bound, or unbound. The case E<0 corresponds to a particle which is confined (and whose energy is less than the well depth) and hence is in a bound state [1]. When E>0, the particle is unconfined and corresponds to a scattering problem. The latter case will only briefly be discussed.

Figure 2.2 [7] provides a prelude to what the wavefunctions and probability distributions for several states will look like in a finite well. Comparing Figure 2.2 with Figure 1.2 for the infinite case, we see that in the finite case, the wavefunctions do not have to be zero at the walls of the well. In the finite case, the wavelengths are slightly longer, implying that the allowed energies will be somewhat smaller.

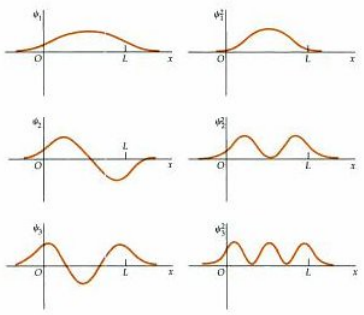


Figure 2.2 In this figure, the points *O* and *L* represent the walls of the well.

Case 1: Bound State

In the bound state we have,−V0<=E<0 (since E cannot be lower than the absolute minimum of the potential [1]. The Schrodinger equation for the two regions is given by [6]

$|x| < L$ (inside the well): $$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) - V_0\psi(x) = -|E|\psi(x)$$

$|x| > L$ (outside the well): $$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) = -|E|\psi(x)$$

Here, the binding energy, $|E|$ of the particle is introduced (|E|=−E). To simplify the TISE equations, let,

$$\alpha = \sqrt{\frac{2m}{\hbar^2}(V_0 - |E|)} \quad \text{and} \quad \beta = \sqrt{\frac{2m}{\hbar^2}|E|}$$

The solutions of the TISE separate into even and odd parity states, and we need only consider positive values of x (which could be inferred from the potential). With the application of the boundary conditions, the even solutions are given as [6]

$$\psi^E(x) = \begin{cases} A\cos(\alpha x) & 0 < x < L \\ Ce^{-\beta x} & x > L \end{cases}$$

and the odd solutions are given as [1]

$$\psi^O(x) = \begin{cases} A\sin(\alpha x) & 0 < x < L \\ Ce^{-\beta x} & x > L \end{cases}$$

Despite the discontinuous nature of the potential at x=L, the wavefunction and its derivative are still continuous, and these conditions provide the required boundary conditions to determine the quantized energies [6]. The requirements that ψO=psiE(x) and Missing open brace for superscript? yields the two equations

$$A\cos(\alpha L) = Ce^{-\beta L}$$
$$-\alpha A\sin(\alpha L) = -\beta Ce^{-\beta L}$$

These can be combined to give the even eigenvalue condition (which depends on the energies but not on the constants A and C) [6] is

$$\alpha\tan(\alpha L) = \beta$$

The odd eigenvalue condition is found to be

$$\alpha\cot(\alpha L) = -\beta$$

The energy levels of the bound states are found by solving these transcendental equations, either graphically or numerically [1]. In doing so, it is helpful to change to dimensionless variables. We introduce the dimensionless quantities [1]

$$\xi = \alpha L$$
$$\eta = \beta L$$

Upon substitution these quantities, the transcendental equations become

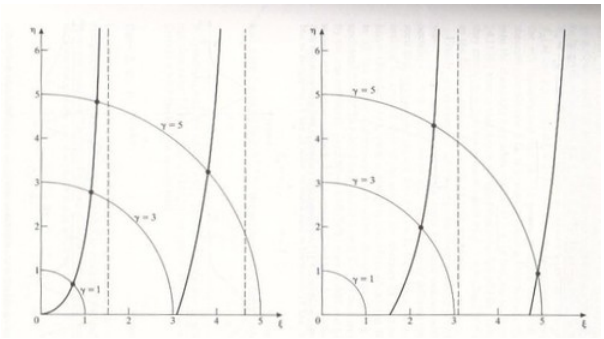
even states: $\xi\tan\xi = \eta$
odd states: $\xi\cot\xi = -\eta$

The graphical determination of the energy levels are obtained by finding the points of intersection of the circle (See Graph 2.1)[1]

$$\xi^2 + \eta^2 = \gamma^2$$

where γ is of known radius (found by simple substitution),

$$\gamma = \sqrt{\frac{2mV_0L^2}{\hbar^2}}$$



Graph 2.1.1 *The left graph shows energy levels for even states, while the right graph represents odd states.*

What can be concluded from this figure?

- The bound-state energy levels are non-degenerate.
- The bound-state energy levels are finite, but increase without bound and depend on the parameter γ. Thus, deeper and wider potentials have a larger number of bound states.
- The bound state spectrum consists of alternating even and odd states, with the ground state always being even.

Case 2: Unbound State

The finite well problem shows that wavefunctions are not localized in the vicinity of the well[10] . This is the case of unbound states, where eigenvalues of energy are a continuum. In this problem, a particle is incident upon the well from the left, interacts with the well, and gets transmitted or reflected. The solution of the Schrodinger equation is outlined briefly. Note that in one dimension, there is a lack of symmetry between the external regions to the left and right of the potential, since the particle is assumed to be incident on the potential in a given direction [6]. Therefore, there is no need to exercise parity.

For the external regions, the solution of the TISE is given by [1]

$$\psi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & x < -L \\ Ce^{ikx} & x > L \end{cases}$$

where

$$k = \sqrt{\frac{2mE}{\hbar^2}}$$

In the region of x<−L, the wavefunction is seen to consist of an incident wave of amplitude *A* and a reflected wave of amplitude *B*. In the region of x>L, the wavefunction is seen as a pure transmitted wave of amplitude *C* [1].

For the internal region, the solution of the TISE is given by [1]

$$\psi(x) = Fe^{i\alpha x} + Ge^{-i\alpha x}$$

where,

$$\alpha = \sqrt{\frac{2m}{\hbar^2}(V_0 + E)}$$

By applying continuity at x=L and x=−L, *F* and *G* are eliminated and the ratios of B/A and C/A can be solved to obtain the reflection coefficient, R=|B/A|2, and the transmission coefficient, T=|C/A|2. An important aspect to briefly point out is that the transmission coefficient is generally less than unity, which is in contradiction to the classical prediction that suggests that the particle should always be transmitted

Summary

To summarize, the major differences between a particle in a finite box and an infinite well, are [((web1))]:

- Only a finite number of energy levels exist (bound state)
- Tunneling into the barrier (wall) is possible
- Higher energy states are less tightly bound than lower ones
- A particle provided with enough energy can escape the well (unbound state)

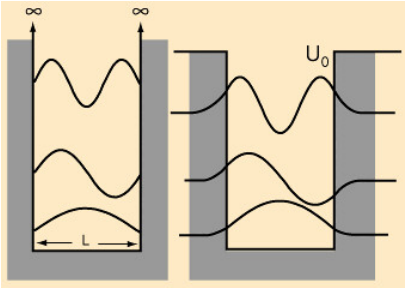


Figure 2.2.1 Comparison of infinite and finite well.

We have considered in some detail a particle trapped between infinitely high walls a distance L apart, we have found the wavefunction solutions of the time independent Schrödinger equation, and the corresponding energies. The essential point was that the wavefunction had to go to zero at the walls, because there is zero probability of finding the particle penetrating an infinitely high wall. This meant that the lowest energy state couldn't have zero energy, that would give a constant nonzero wavefunction. Rather, the lowest energy state had to have the minimal amount of bending of the wavefunction necessary for it to be zero at the two walls but nonzero in between-this corresponds to half a period of a sine or cosine (depending on the choice of origin), these functions being the solutions of Schrödinger's equation in the zero potential region between the walls. The sequence of wavefunctions (eigenstates) as the energy increases have 0, 1, 2, … zeros (nodes) in the well.

Let us now consider how this picture is changed if the potential at the walls is not infinite. It will turn out to be convenient to have the origin at the center of the well, so we take

- $V(x) = V_0$ for $x < -L/2$
- $V(x) = 0$ for $-L/2 < x < L/2$
- $V(x) = V_0$ for $L/2 < x$.

Having the potential symmetric about the origin makes it easier to catalog the wavefunctions. For a symmetric potential, the wavefunctions can always be taken to be symmetric or antisymmetric.

Symmetry

If a wavefunction $\psi(x)$ is a solution of Schrödinger's equation with energy E, and the potential is symmetric, then $\psi(-x)$ is a solution with the same energy. This means that $\psi(x) + \psi(-x)$ and $\psi(x) - \psi(-x)$ are also solutions, since the equation is linear, and these are *symmetric* and *antisymmetric* respectively, and using them is completely equivalent to using the original $\psi(x)$ and its reflection $\psi(-x)$.

How is the lowest energy state wavefunction affected by having finite instead of infinite walls? Inside the well, the solution to Schrödinger's equation is still of cosine form (it's a state symmetric about the origin). However, since the walls are now finite, $\psi(x)$ cannot change slope discontinuously to a flat line at the walls. It must instead connect smoothly with a function which is a solution to Schrödinger's equation *inside the wall*.

The equation in the wall is

$$\frac{-\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V_0\psi(x) = E(x) \qquad (3.9B.1)$$

and has two exponential solutions (say, for \(x > L/2\)) one increasing to the right, the other decreasing,

$$e^{\alpha x} \text{ and } e^{-\alpha x}$$

where

$$\alpha = \sqrt{2m(V_o - E)/hbar^2} \tag{3.9B.2}$$

(We are assuming here that $E < V_0$, so the particle is bound to the well. We shall find this is always true for the lowest energy state.)

Let us try to construct the wavefunction for the energy E corresponding to this lowest bound state. From the equation with $V_0 = 0$, the wavefunction inside the well (let's assume it's symmetric for now) is proportional to $\cos kx$, where $k = \sqrt{2mE/\hbar^2}$.

The wavefunction (and its derivative!) inside the well must match a sum of exponential terms—the wavefunction in the *wall*—at \($x = L/2$\), so

$$\cos\left(\frac{kL}{2}\right) = Ae^{\alpha L/2} + Be^{-\alpha L/2} \tag{3.9B.3}$$

$$-k\sin\left(\frac{kL}{2}\right) = \alpha e^{\alpha L/2} - \alpha e^{-\alpha L/2} \tag{3.9B.4}$$

(By writing just a cosine term inside the well, we have left out the overall normalization constant. This can be put back in at the end.)

Solving these equations for the coefficients A, B in the usual way, we find that in general the cosine solution inside the well goes smoothly into a linear combination of exponentially increasing and decreasing terms in the wall. (By the symmetry of the problem, the same thing must happen for x < -L/2.) *However, this cannot in general represent a bound state in the well*. The increasing solution increases *without limit* as \(x\) goes to infinity, so since the square of the wavefunction is proportional to the probability of finding the particle at any point, the particle is infinitely more likely to be found at infinity than anywhere else. It got away! This clearly makes no sense—we're trying to find wavefunctions for particles that stay in, or at least close to, the well. We are forced to conclude that the *only* exponential wavefunction that makes sense is the one for which *A is exactly zero*, so that there is only a *decreasing* wave in the wall.

Requiring the decreasing wavefunction, \($A = 0$\), means that only a discrete set of values of k, or E, satisfy the boundary condition equations above. They are most simply found by taking \($A = 0$\) and dividing one equation by the other to give:

$$\tan\left(\frac{kL}{2}\right) = \frac{\alpha}{k} \tag{3.9B.5}$$

This cannot be solved analytically, but is easy to solve graphically by plotting the two sides as functions of \(k\) and finding where the curves intersect.

Using the Spreadsheet for the Finite Square Well

Recall that by taking the origin in the center of the square well, we argued that we need only look at wavefunctions that were symmetric or antisymmetric about the origin. If you try to sketch such wavefunctions, you will find that symmetric wavefunctions must have zero slope at the origin, and antisymmetric wavefunctions must be zero at the origin.

Furthermore, if we know the wavefunction in the right-hand half, that is, for x > 0, we know it for all x, from the symmetry. Hence, we need only solve Schrödinger's equation going from the origin to the right - we can take x_0 = 0, x_1 = dx, etc.

As we discussed in the preceding section, on integrating out from the origin with energy E set less than V_0, once we cross over into the wall the solution is some combination of an exponentially increasing function and an exponentially decreasing function,

$$f(x) = A(E)e^{\alpha x} + B(E)e^{-\alpha x}$$

where the exponential coefficient α is positive, and depends on E—it was defined in the preceding section.

Let us first look at the *symmetric* solutions for very low energies, so take f(0) = 1, f'(0) =0. (Note that we cannot find the correct overall *normalization constant* until we find the solution, then integrate its square over all space—this can always be done later, and is unnecessary for analyzing the properties of the state).

Let us begin with the trivial case E = 0. For zero energy, inside the well E - V will of course be identically zero, so from Schrödinger's equation the *slope* of $f(x)$ can never change, consequently $f(x)$ = 1 for all x < L. On reaching the wall, this wavefunction and its derivative connect smoothly from inside to outside if A = B = ½. It is clear that as we keep going to the right, the A term in the above equation dominates and $f(x)$ diverges, signaling that there is no state localized in the well at E = 0.

At this point, you should create an Excel spreadsheet, with a graph showing $f(x)$ and $V(x)$-E as functions of x. Details of how this might be done are given in the accompanying homework assignment. *If you have a working spreadsheet, it will be **much** easier to understand the following discussion!*

As we now increase E from zero, the symmetric wavefunction, having zero slope at the center of the well, will begin to curve downwards on moving away from the center, and as the energy increases so does the downward curvature. This naturally changes the mix of increasing and decreasing exponentials needed to connect smoothly at the wall, and in fact as a function of E, $A(E)$ changes sign at a certain value we will call E_0. For energies just below E_0, $f(x)$ diverges to plus infinity for large x. For energies above E_0, it diverges to minus infinity for large x. *Exactly at E_0, $f(x)$ goes to zero for large x.* This is the wavefunction we are looking for: it corresponds to a particle localized close to the well, and in fact is the lowest possible energy—the *ground state*—for a particle in the well. E_0 is called the ground state *eigenvalue*, the wavefunction is called an *eigenstate*.

Finding the Ground State Energy

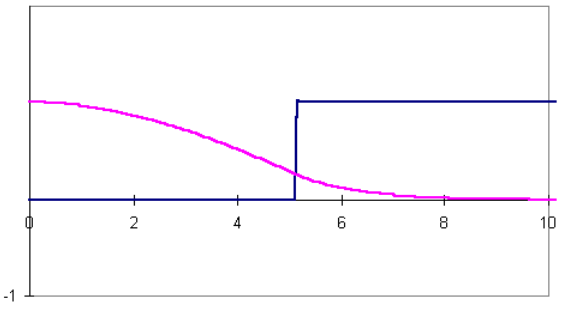
The first problem is to find this value E_0. If we just guess a value of E, the wavefunction will almost certainly diverge for large x. The way to find E_0 is to notice that for E below E_0, $f(x)$ goes to large positive values on the far right, for E just above E_0, $f(x)$ goes to large negative values on the far right. So we take an increasing set of E values starting near zero, and watch for the tail to wag! When this happens, we back up half way, then back or forward as necessary, choosing a set of E values that bracket E_0.

Actually, this rather tedious process can be automated in Excel, but it's worth going through it manually a time or two to get a feeling for how the wavefunctions behave. A general feature to check is that for $E > V(x)$ they are oscillatory, sine or cosine like, although with changing wavelengths in general; whereas for $E < V(x)$ they have exponential character. You can see why this is by looking at the equation, but it’s worth bearing in mind as you examine the curves.

Automating the Process

Now to a bit of automation. The strategy we are following is to adjust E until the function $f(x)$ goes to zero for large x, that is, far down at the bottom of our column having several hundred entries. Since we would like to know how well we are doing, it’s convenient to have the spreadsheet copy that value to some cell near the top. So, if the end of our column of $f(x_i)$'s is F600, say, then we enter =F600 in some convenient square near the top, say B22. Now we can keep an eye on the distant value as we adjust E. We can also get Excel to do this for us as follows (I assume the value of E is entered into square B15):

The curve below was generated by this procedure. It is worth noting that there is a definite probability of finding the particle inside the wall, that is to say, the wavefunction is nonzero there (and, it turns out, for reasonable, physical values of the parameters):

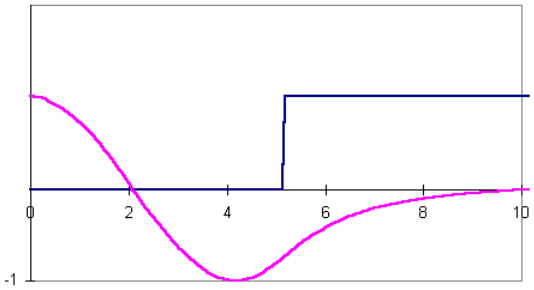


All square well potentials in one dimension, however shallow, have a localized ground state with this general shape. Whether or not there are other eigenstates with other eigenvalues depends on the depth of the potential. For a sufficiently shallow potential, there is only one state. An infinitely deep well, as we discussed earlier, has an infinite number of bound states. As the well depth increases from zero, states are bound sequentially. These higher eigenstates are called *excited states*. A particle in one of them will usually decay to a lower state, emitting a photon, just as in the Bohr atom.

https://chem.libretexts.org/@go/page/198662

Finding Excited States

Let us now see how the wavefunction develops as E increases beyond E_0 for a sufficiently deep well. As the energy is increased, the cosine term inside the well has tighter curvature, and the exponentially increasing term for large x is large and negative. However, continuing to increase the energy, eventually the cosine term bottoms out inside the well and begins to turn up again before reaching the wall. At a certain energy, it again becomes possible to match it to a decreasing function:



To find this new eigenvalue, we can get close to it then use: Tools, Goal Seek just as we did for the ground state—see above.

A point to notice about the wavefunction pictured above is that has a node at about $x = 2$. (Actually, of course, this is just *half* the wavefunction for the complete well-there will be another node at $x = -2$.) As we continue to increase E, the high-x tail of the wavefunction wags up and down. Each time it crosses the axis there is an allowable wavefunction. Furthermore, each time it crosses the axis the wavefunction collects another node for $x > 0$. Thus, the complete wavefunctions generated by this method have 0, 2, 4, 6, … nodes. From the symmetry of the problem, the allowable wavefunctions with an *odd* number of nodes must have one node at the origin. They can be generated by taking as initial conditions that f is zero at the origin, and has finite slope.

Curvature of wavefunctions

Schrödinger's equation in the form

$$\frac{d^2 f(x)}{dx^2} = \{V(x) - E\} f(x)$$

can be interpreted by saying that the left-hand side, the rate of change of slope, is the *curvature*—so the curvature of the function is equal to $(V(x) - E)f(x)$. This means that if $E > V(x)$, for $f(x)$ positive $f(x)$ is curving negatively, for $f(x)$ negative $f(x)$ is curving positively. *In both cases, $f(x)$ is always curving towards the axis*. This means that for $E > V(x)$, $f(x)$ has a kind of stability: its curvature is always bringing it back towards the axis, so it has oscillatory character.

On the other hand, for $V(x) > E$, the curvature is always *away* from the axis. This means that $f(x)$ tends to diverge to infinity. Only under exactly the right conditions will this curvature be just enough to bring the wavefunction to zero as x goes to infinity. (As $f(x)$ tends to zero, the curvature tends to zero, too.)

It is worth examining the wavefunctions generated by the spreadsheet to see just how the curvature changes as $V(x) - E$, or for that matter $f(x)$, changes sign.

Tunneling

https://chem.libretexts.org/Core/Phy...nics/Tunneling

Contributors

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

3.9B: Particle in a Finite Box and Tunneling (optional) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.9: The Uncertainty Principle Redux - Estimating Uncertainties from Wavefunctions

Learning Objectives

- Expand on the introduction of Heisenberg's Uncertainty Principle by calculating the Δx or Δp directly from the wavefunction

As will be discussed in Section 4.6, the operators $\hat{x}$ and $\hat{p}$ are not compatible and there is **no** measurement that can precisely determine the coresponding observables (x and p) simultaneously. Hence, there must be an uncertainty relation between them that specifies how uncertain we are about one quantity given a definite precision in the measurement of the other. Presumably, if one can be determined with infinite precision, then there will be an infinite uncertainty in the other. The uncertainty in a general quantity A is

$$\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2} \tag{3.9.1}$$

where $\langle A^2 \rangle$ and $\langle A \rangle$ are the expectation values of $\hat{A}^2$ and $\hat{A}$ operators for a specific wavefunction. Extending Equation 3.9.1 to x and p results in the following uncertainties

$$\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} \tag{3.9.2}$$

$$\Delta p = \sqrt{\langle p^2 \rangle - \langle p \rangle^2} \tag{3.9.3}$$

These quantities can be expressed explicitly in terms of the (time-dependent) wavefunction $\Psi(x,t)$ using the fact that

$$\langle x \rangle = \langle \Psi(t) \left| \hat{x} \right| \Psi(t) \rangle \tag{3.9.4}$$

$$= \int \Psi^*(x,t) x \Psi(x,t)\, dx \tag{3.9.5}$$

and

$$\langle x^2 \rangle = \langle \Psi(t) \left| \hat{x}^2 \right| \Psi(t) \rangle \tag{3.9.6}$$

$$= \int \Psi^*(x,t) x^2 \Psi(x,t)\, dx \tag{3.9.7}$$

The middle terms in both Equations 3.9.4 and 3.9.6 are the integrals expressed in Dirac's Bra-ket notation. Similarly using the definition of the linear momentum operator:

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x}.$$

So

$$\langle p \rangle = \langle \Psi(t) \left| \hat{p} \right| \Psi(t) \rangle \tag{3.9.8}$$

$$= \int \Psi^*(x,t) - i\hbar \frac{\partial}{\partial x} \Psi(x,t)\, dx \tag{3.9.9}$$

and

$$\langle p^2 \rangle = \langle \Psi(t) \left| \hat{p}^2 \right| \Psi(t) \rangle \tag{3.9.10}$$

$$= \int \Psi^*(x,t) \left(-\hbar^2 \frac{\partial^2}{\partial x^2} \right) \Psi(x,t)\, dx \tag{3.9.11}$$

Time-dependent vs. time-independent wavefunction

The expectation values above are formulated with the total time-dependence wavefunction $\psi(x,t)$ that are functions of x and t. However, it is easy to show that the same expectation value would be obtained if the time-independent wavefunction $\psi(x)$ that are functions of only x are used. If $V(x)$ in $\hat{H}$ is time independent, then the wavefunctions are stationary and the expectation

value are time-independent. You can easily confirm that by comparing the expectation values using the general formula for a stationary wavefunction

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

and for $\psi(x)$.

The Heisenberg uncertainty principle can be quantitatively connected to the properties of a wavefunction, i.e., calculated via the expectation values outlined above:

$$\Delta p \Delta x \geq \frac{\hbar}{2} \tag{3.9.12}$$

This essentially states that the greater certainty that a measurement of x or p can be made, the greater will be the *uncertainty* in the other. Hence, as Δp approaches 0, Δx must approach ∞, which is the case of the free particle (e.g, with $V(x) = 0$) where the momentum of a particles can be determined precisely.

Example 3.8.1 : Uncertainty with a Gaussian wavefunction

A particle is in a state described by the wavefunction

$$\psi(x) = \left(\frac{2a}{\pi}\right)^{\frac{1}{4}} e^{-ax^2} \tag{3.9.13}$$

where a is a constant and $-\infty \leq x \leq \infty$. Verify that the value of the product $\Delta p \Delta x$ is consistent with the predictions from the uncertainty principle (Equation 3.9.12).

Solution

Let's calculate the average of x:

$$\begin{aligned}\langle x \rangle &= \int_{-\infty}^{\infty} \psi^* x \psi \, dx \\ &= \int_{-\infty}^{\infty} (2a/\pi)^{\frac{1}{4}} e^{-ax^2} x (2a/\pi)^{\frac{1}{4}} e^{-ax^2} \, dx \\ &= \int_{-\infty}^{\infty} x (2a/\pi)^{\frac{1}{2}} e^{-2ax^2} \, dx \\ &= 0\end{aligned}$$

since the integrand is an odd function (an even function times an odd function is an odd function). This makes sense given that the gaussian wavefunction is symmetric around $x = 0$.

Let's calculate the average of x^2:

$$\begin{aligned}\langle x^2 \rangle &= \int_{-\infty}^{\infty} \psi^* x^2 \psi \, dx \\ &= \int_{-\infty}^{\infty} (2a/\pi)^{\frac{1}{4}} e^{-ax^2} (x^2) (2a/\pi)^{\frac{1}{4}} e^{-ax^2} \, dx \\ &= \int_{-\infty}^{\infty} x^2 (2a/\pi)^{\frac{1}{2}} e^{-2ax^2} \, dx \\ &= \frac{1}{4a}\end{aligned}$$

Let's calculate the average in p:

$$\begin{aligned}\langle p \rangle &= \int_{-\infty}^{\infty} \psi^* p \psi \, dx \\ &= \int_{-\infty}^{\infty} \left(\frac{2a}{\pi}\right)^{\frac{1}{4}} e^{-ax^2} - i\hbar \frac{d}{dx} \left(\frac{2a}{\pi}\right)^{\frac{1}{4}} e^{-ax^2} \, dx \\ &= \int_{-\infty}^{\infty} \left(\frac{2a}{\pi}\right)^{\frac{1}{4}} e^{-ax^2} (-i\hbar) \left(\frac{2a}{\pi}\right)^{\frac{1}{4}} e^{-ax^2} (-2ax) \, dx \\ &= 0\end{aligned}$$

since the integrand is an odd function.

Let's calculate the average of p^2:

$$\begin{aligned}\langle p^2\rangle &= \int_{-\infty}^{\infty}\psi^* p^2\psi dx \\ &= -\hbar^2\left(\frac{2a}{\pi}\right)^{1/2}\int_{-\infty}^{\infty}2a(ax^2-1)e^{-2ax^2}dx \\ &= -4\hbar^2a^2\left(\frac{2a}{\pi}\right)^{1/2}\int_0^{\infty}x^2e^{-2ax^2}dx + 4\hbar^2a\left(\frac{2a}{\pi}\right)^{1/2}\int_0^{\infty}e^{-2ax^2}dx \\ &= a\hbar^2\end{aligned}$$

We use Equation 3.9.1 to check on the uncertainty

$$\Delta x^2 = \langle x^2\rangle - \langle x\rangle^2 = \frac{1}{4a} - 0$$

$$\Delta x = \sqrt{\Delta x^2} = \frac{1}{2\sqrt{a}}$$

$$\Delta p^2 = \langle p^2\rangle - \langle p\rangle^2 = a\hbar^2 - 0$$

$$\Delta p = \sqrt{\Delta p^2} = \hbar\sqrt{a}$$

Finally we have

$$\Delta p\Delta x = \left(\frac{1}{2\sqrt{a}}\right)(\hbar\sqrt{a}) = \frac{\hbar}{2}$$

Not only does the Heisenburg uncertainly principle hold (Equation 3.9.12), but the equality is established for this wavefunction. This is because the Gaussian wavefunction (Equation 3.9.13) is special as discussed later.

? Exercise 3.8.1

A particle is in a state described by the ground state wavefunction of a particle in a box

$$\psi = \sqrt{\frac{2}{L}}\sin\left(\frac{\pi x}{L}\right)$$

where L is the length of the box and $0 \le x \le L$. Verify that the value of the product $\Delta p\Delta x$ is consistent with the predictions from the uncertainty principle (Equation 3.9.12).

The uncertainty principle is a consequence of the wave property of matter. A wave has some finite extent in space and generally is not localized at a point. Consequently there usually is significant uncertainty in the position of a quantum particle in space.

3.9: The Uncertainty Principle Redux - Estimating Uncertainties from Wavefunctions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.8: The Uncertainty Principle - Estimating Uncertainties from Wavefunctions** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

3.10: A Particle in a Two-Dimensional Box

Learning Objectives

- Manipulate probability, energy, and transition problems involving multidimensional wavefunctions
- Apply the Separation of Variables to solve a multi-dimensional Schrödinger equation
- Manipulate problems with multiple quantum numbers
- To be introduced to the concepts of non-singular nodes (e.g., nodal lines)
- To be introduced to the concept of degeneracy

Let us now consider the Schrödinger Equation for an electron confined to a two dimensional box, $0 < x < a$ and $0 < y < b$. That is to say, within this rectangle the electron wavefunction behaves as a free particle ($V(x,y) = 0$), but the walls are impenetrable so the wavefunction $\psi(x,y,t) = 0$ at the walls.

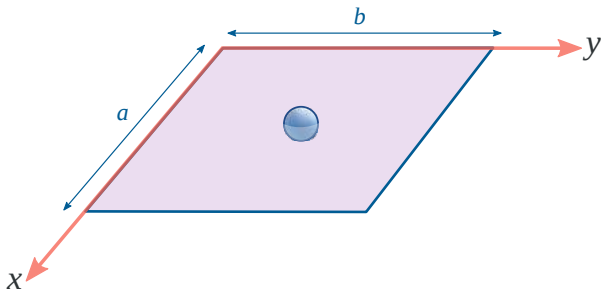


Figure 3.10.1: A particle in a $a \times b$ 2-dimensional box with infinite height can only be found within the box (highlighted). (CC BY-NC; Ümit Kaya)

Extending the (time-independent) Schrödinger equation for a one-dimensional system

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \tag{3.10.1}$$

to a two-dimensional system is not difficult

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi(x,y)}{\partial x^2} + \frac{\partial^2\psi(x,y)}{\partial y^2}\right) + V(x,y)\psi(x,y) = E\psi(x,y). \tag{3.10.2}$$

Equation 3.10.2 can be simplified for the particle in a 2D box since we know that $V(x,y) = 0$ within the box and $V(x,y) = \infty$ outside the box, ii.e.

$$V(x,y) = \begin{cases} 0 & 0 \le x \le a \text{ and } 0 \le y \le b \\ \infty & x < 0 \text{ and } x > a \\ \infty & y < 0 \text{ and } y > b \end{cases}$$

So Equation 3.10.2 becomes

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi(x,y)}{\partial x^2} + \frac{\partial^2\psi(x,y)}{\partial y^2}\right) = E\psi(x,y). \tag{3.10.3}$$

Since the Hamiltonian (i.e. left side of Equation 3.10.3) is the sum of two terms with independent (separate) variables, we try a product wavefunction like in the Separation of Variables approach used to separate time-dependence from the spatial dependence previously. Within this approach we express the 2-D wavefunction as a product of two independent 1-D components

$$\psi(x,y) = X(x)Y(y). \tag{3.10.4}$$

This ansatz separates Equation 3.10.3 into two independent one-dimensional Schrödinger equations

$$-\frac{\hbar^2}{2m}\left(\frac{d^2X(x)}{dx^2}\right) = \varepsilon_x X(x). \tag{3.10.5}$$

$$-\frac{\hbar^2}{2m}\left(\frac{d^2Y(y)}{dy^2}\right) = \varepsilon_y Y(y). \tag{3.10.6}$$

where the total energy of the particle is the sum of the energies from each one-dimensional Schrödinger equation

$$E = \varepsilon_x + \varepsilon_y \tag{3.10.7}$$

The differential equations in Equations 3.10.5 and 3.10.6 are familiar as they were found for the particle in a 1-D box previously. They have the general solution

$$X(x) = A_x \sin(k_x x) + B_x \cos(k_x x) \tag{3.10.8}$$

$$Y(y) = A_y \sin(k_y y) + B_y \cos(k_y y) \tag{3.10.9}$$

Applying Boundary Conditions

The general solutions in Equations 3.10.8 and 3.10.9 can be simplified to address boundary conditions dictated by the potential, i.e., $\psi(0,y) = 0$ and $\psi(x,0) = 0$. Therefore, $B_x = 0$ and $B_y = 0$.

Thus, we can combine Equations 3.10.8, 3.10.9, and 3.10.4 to construct the wavefunction $\psi(x,y)$ for a particle in a 2D box of the form

$$\psi(x,y) = N \sin\left(\sqrt{\frac{2m\varepsilon_x}{\hbar^2}}x\right) \sin\left(\sqrt{\frac{2m\varepsilon_y}{\hbar^2}}y\right) \tag{3.10.10}$$

We still need to satisfy the remaining boundary conditions $\psi(L,y) = 0$ and $\psi(x,L) = 0$

$$N \sin\left(\sqrt{\frac{2m\varepsilon_x}{\hbar^2}}L\right) \sin\left(\sqrt{\frac{2m\varepsilon_y}{\hbar^2}}y\right) = 0 \tag{3.10.11}$$

and

$$N \sin\left(\sqrt{\frac{2m\varepsilon_x}{\hbar^2}}x\right) \sin\left(\sqrt{\frac{2m\varepsilon_y}{\hbar^2}}L\right) = 0 \tag{3.10.12}$$

Equation 3.10.11 can be satisfied if

$$\sin\left(\sqrt{\frac{2m\varepsilon_x}{\hbar^2}}L\right) = 0 \tag{3.10.13}$$

independent of the value of y, while Equation 3.10.12 can be satisfied if

$$\sin\left(\sqrt{\frac{2m\varepsilon_y}{\hbar^2}}L\right) = 0 \tag{3.10.14}$$

independent of the value of x. These are the same conditions that we encountered for the one-dimensional box, hence we already know the $\sin$ function in each case can be zero in many places. In fact, these two conditions are satisfied if

$$\sqrt{\frac{2m\varepsilon_x}{\hbar^2}}L = n_x\pi \tag{3.10.15}$$

and

$$\sqrt{\frac{2m\varepsilon_y}{\hbar^2}}L = n_y\pi \tag{3.10.16}$$

which yield the allowed values of ε_x and ε_y as

$$\varepsilon_{n_x} = \frac{\hbar^2\pi^2}{2mL^2}n_x^2 \tag{3.10.17}$$

and

$$\varepsilon_{n_y} = \frac{\hbar^2\pi^2}{2mL^2}n_y^2 \tag{3.10.18}$$

We need two different integers n_x and n_y because the conditions are completely independent and can be satisfied by any two different (or similar) values of these integers. The allowed values of the total energy are now given by

$$E_{n_x,n_y} = \frac{\hbar^2\pi^2}{2mL^2}(n_x^2 + n_y^2) \tag{3.10.19}$$

Note that the allowed energies now depend on two integers n_x and n_y rather than one. These arise from the *two* independent boundary conditions in the x and y directions. As in the one-dimensional box, the values of n_x and n_y are both restricted to the natural numbers $1, 2, 3, ...$ Note, therefore, that the ground state energy $E_{1,1}$ is

$$E_{1,1} = \frac{\hbar^2\pi^2}{mL^2} \tag{3.10.20}$$

is larger than for the one-dimensional box because of the contributions from kinetic energy in the x and y directions.

Once the conditions on ε_{n_x} and ε_{n_y} are substituted into Equation 3.10.10, the wavefunctions become

$$\psi_{n_x,n_y}(x,y) = N\sin\left(\frac{n_x\pi x}{L}\right)\sin\left(\frac{n_y\pi y}{L}\right) \tag{3.10.21}$$

The constant N is now determined by the normalization condition

$$\begin{aligned}\int_0^L\int_0^L |\psi_{n_x,n_y}(x,y)|^2 dx\,dy &= 1\\ N^2\int_0^L \sin^2\left(\frac{n_x\pi x}{L}\right)dx \int_0^L \sin^2\left(\frac{n_y\pi y}{L}\right)dy &= 1\\ N^2\frac{L}{2}\cdot\frac{L}{2} &= 1\\ N &= \frac{2}{L}\end{aligned}$$

so that the complete normalized 2D wavefunction is

$$\psi_{n_x,n_y}(x,y) = \frac{2}{L}\sin\left(\frac{n_x\pi x}{L}\right)\sin\left(\frac{n_y\pi y}{L}\right) \tag{3.10.22}$$

The Separability of the Hamiltonian

Because the Halmiltonian in Equation 3.10.3 can be expressed as a simple sum of independent terms, the solutions can be expressed as simple **products** of the one-dimensional solution (Equation 3.10.22) with energies that are expressed as the **sum** of one-dimensional energies (Equation 3.10.19). This observation extends to higher dimensional wavefunctions as demonstrated later.

The wavefunctions in Equation 3.10.22 are somewhat more difficult that 1-D analogs to visualize because they are two dimensional. Nevertheless, we can still visualize them, and Figure 3.10.1 shows the following wavefunctions: $\psi_{1,1}(x,y)$, $\psi_{2,1}(x,y)$ and $\psi_{2,2}(x,y)$.

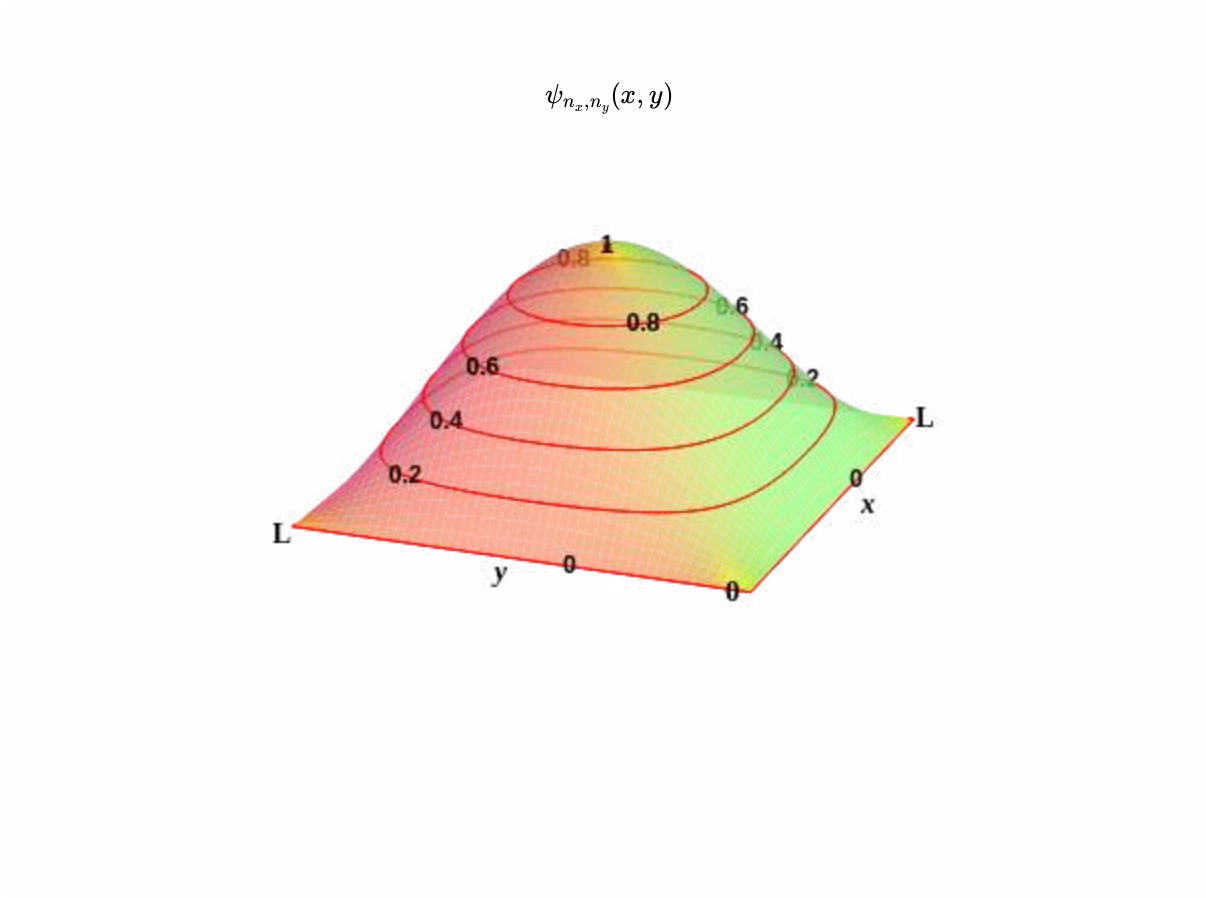


Figure 3.10.1*: Visualizing the first six wavefunctions for a particle in a two-dimensional square box ($L_x = L_y = L$). Use the slide bar to independently change either n_x or n_y quantum number and see the changing wavefunction.*

It is important to note that the wavefunctions can be either positive or negative, even though the associated probability density $p_{n_x,n_y}(x,y) = |\psi_{n_x,n_y}(x,y)|^2$ is strictly positive.

https://chem.libretexts.org/@go/page/198664

$\psi_{n_x,n_y}(x, y)$ $|\psi_{n_x,n_y}(x, y)|^2$

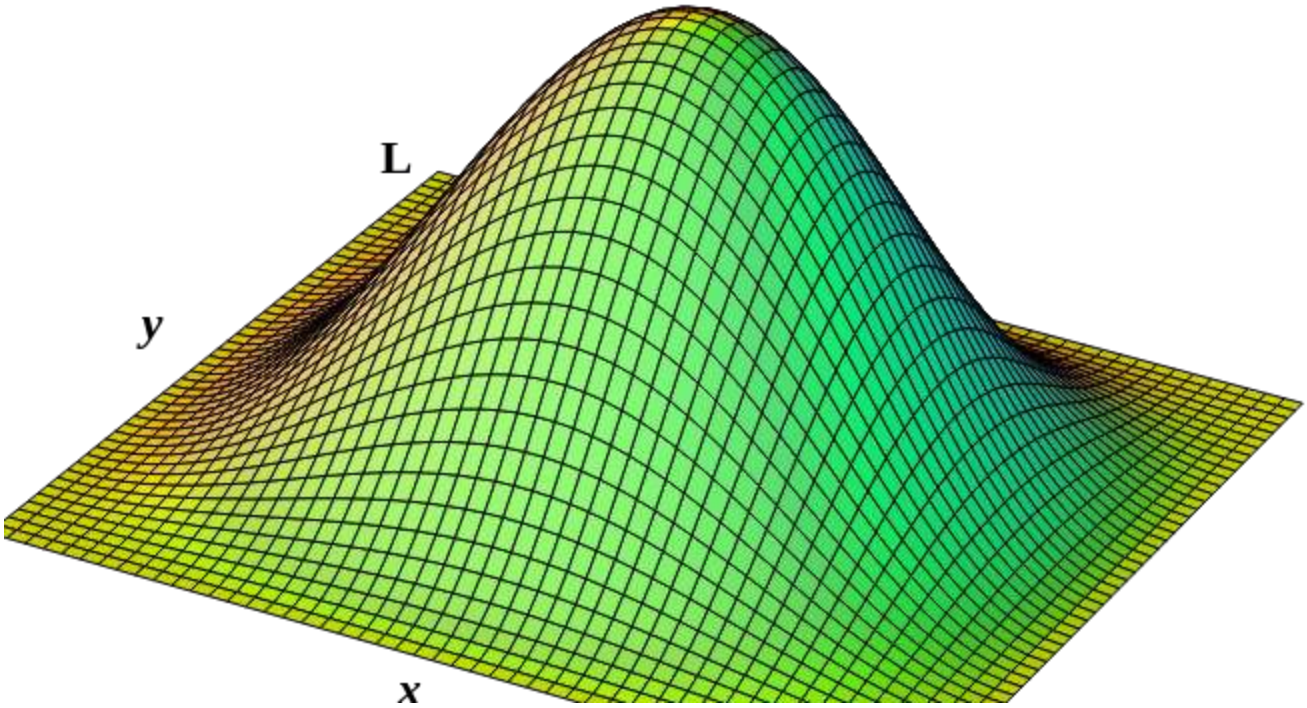


Figure **3.10.2**: *Visualizing the first six wavefunctions and associated probability densities for a particle in a two-dimensional square box ($L_x = L_y = L$). Use the slide bar to independently change either n_x or n_y quantum number and see the changing wavefunction.*

Unlike in the one-dimensional analoge, where nodes in the wavefunction are points where $\psi_n(x) = 0$, here entire lines can be nodal (called *nodal lines)*. For example, in the state $\psi_{2,1}(x, y)$, there is a nodal line at $\psi_{2,1}(L/2, y)$. Along the entire line $x = L/2$, the wavefunction is 0 independent of the value of y. The wavefunction $\psi_{2,2}(x, y)$ has two nodal lines when $x = L/2$ and when $y = L/2$. The sign of the wavefunction and its nodal structure will play central roles later when we consider chemical bonding.

The fact that the wavefunction $\psi_{n_x n_y}(x, y)$ is a product of one-dimensional wavefunctions:

$$\psi_{n_x n_y}(x, y) = \psi_{n_x}(x)\psi_{n_y}(y) \tag{3.10.23}$$

makes the calculation of probabilities rather easy. The probability that a measurement the particle's position will yield a value of $x \in [a, b]$ and $y \in [c, d]$ is

$$\begin{aligned} P(x \in [a, b] \text{ and } y \in [c, d]) &= \int_a^b dx \int_c^d dy |\psi_{n_x n_y}(x, y)|^2 \\ &= \int_a^b dx \int_c^d dy \psi_{n_x}^2(x)\psi_{n_y}^2(y) \\ &= \left[\int_a^b \psi_{n_x}^2(x)dx\right]\left[\int_c^d \psi_{n_y}^2(y)dy\right] \end{aligned}$$

Example 3.10.1: Probability

For a particle in a two-dimensional square box of length L, if the particle is in the $\psi_{1,2}(x,y)$ state, what is the probability that a measurement of the particle's position will yield $x \in [0, L/2]$ and $y \in [0, L/2]$?

Solution

Substituting the limits of integration into Equation ???, we have

$$\begin{aligned} P(x \in [0, L/2] \text{ and } y \in [0, L/2] &= \left[\int_0^{L/2} \psi_1^2(x)dx\right]\left[\int_0^{L/2} \psi_2^2(y)dy\right] \\ &= \left[\frac{2}{L}\int_0^{L/2} \sin^2\left(\frac{\pi x}{L}\right)dx\right]\left[\frac{2}{L}\int_0^{L/2} \sin^2\left(\frac{2\pi y}{L}\right)dy\right] \\ &= \left[\frac{2}{L}\frac{L}{4}\right]\left[\frac{2}{L}\frac{L}{4}\right] \\ &= \frac{1}{4} \end{aligned}$$

This can be graphically confirmed by exploring the probability densities in Figure **3.10.2** for the $\psi_{1,2}(x,y)$ state.

Exercise 3.10.1

For a particle in a two-dimensional square box of length L, if the particle is in the state $\psi_{2,3}(x,y)$, what is the probability that a measurement of the particle's position will yield $x \in [0, L/2]$ and $y \in [0, L/3]$?

Note that if the box were rectangular rather than square, then instead of having a length of L on both sides, there would be two different lengths L_x and L_y. The formulas for the energies and wavefunctions become only slightly more complicated:

$$E_{n_x,n_y} = \frac{\hbar^2\pi^2}{2m}\left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2}\right) \tag{3.10.24}$$

instead of Equation 3.10.19 and

$$\psi_{n_x,n_y}(x,y) = \frac{2}{\sqrt{L_x L_y}}\sin\left(\frac{n_x\pi x}{L_x}\right)\sin\left(\frac{n_y\pi y}{L_y}\right) \tag{3.10.25}$$

instead of Equation 3.10.22.

Degeneracy

Two distinct wavefunctions are said to be *degenerate* if they correspond to the same energy. If the sides *a*, *b* of the rectangle are such that *a*/*b* is irrational (the general case), there will be no degeneracies. The *most* degenerate case is the square, $L_x = L_y$, for which clearly $E_{m,n} = E_{n,m}$. Degeneracies in quantum physics are most often associated with symmetries in this way. Figure **3.10.3** shows the wavefunctions (3,2) and (2,3) for a rectangle. These are contour maps for the time-independent solution, with white being the highest point. These two wavefunctions do not correspond to the same energy, although they would, of course, for a square.

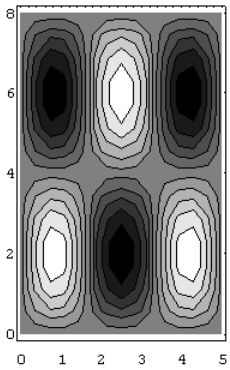


Figure 3.10.2: Contours of the $n_x = 2$ and $n_y = 3$ wavefunction (left) and $n_x = 3$ and $n_y = 2$ wavefunction (right).

The energy of the particle in a 2-D square box (i.e., $L_x = L_y = L$) in the **ground state** is given by Equation 3.10.24 with $n_x = 1$ and $n_y = 1$. This energy (E_{11}) is hence

$$E_{1,1} = \frac{2\hbar^2\pi^2}{2mL^2} \tag{3.10.26}$$

For the ground state of the particle in a 2D box, there is one wavefunction (and no other) with this specific energy; the ground state and the energy level are said to be **non-degenerate**. However, in the 2-D box potential, the energy of a state depends upon the sum of the squares of the two quantum numbers. The particle having a particular value of energy in the excited state *may* has several different stationary states or wavefunctions. If so, these states and energy eigenvalues are said to be **degenerate**.

For the first excited state, three combinations of the quantum numbers $(n_x,\ n_y)$ are $(2,\ 1)$ and $(1, 2)$. The sum of squares of the quantum numbers in each combination is same (equal to 5). Each waveunction has same energy:

$$E_{2,1} = E_{1,2} = \frac{5\hbar^2\pi^2}{2mL^2} \tag{3.10.27}$$

Corresponding to these combinations three different wavefunctions and **two** different states are possible. Hence, the first excited state is said to be doubly degenerate. The number of independent wavefunctions for the stationary states with *a shared energy* is called as the **degree of degeneracy** of the energy level. The value of energy levels with the corresponding combinations and sum of squares of the quantum numbers

$$n^2 = n_x^2 + n_y^2 \tag{3.10.28}$$

as well as the degree of degeneracy are depicted in Table 3.10.1.

Table 3.10.1: Degeneracy properties of the particle in a 2-D box with $L_x = L$ and $L_y = L$.

$n_x^2 + n_y^2$	Combinations of Degeneracy $(n_x,\ n_y)$			Total Energy (E_{n_x,n_y})	Degree of Degeneracy
2	(1, 1)			$\frac{2\hbar^2\pi^2}{2mL^2}$	1
5	(2, 1)	(1, 2)		$\frac{5\hbar^2\pi^2}{2mL^2}$	2
8	(2, 2)			$\frac{8\hbar^2\pi^2}{2mL^2}$	1
10	(3, 1)	(1, 3)		$\frac{10\hbar^2\pi^2}{2mL^2}$	2
13	(3, 2)	(2, 3)		$\frac{13\hbar^2\pi^2}{2mL^2}$	2
18	(3,3)			$\frac{18\hbar^2\pi^2}{2mL^2}$	1

Note that all the states above would be non-degenerate if $L_x \neq L_y$. However, a degeneracy will be "accidentally" observed for states that fulfill this requirement:

$$\frac{n_x^2}{L_x^2} = \frac{n_y^2}{L_y^2}. \tag{3.10.29}$$

Example 3.10.2: Transition Energy

An electron in a 2D infinite potential well needs to absorb electromagnetic wave with wavelength 4040 nm (IR radiation) to be excited from lowest excited state to next higher energy state. What is the length of the box if this potential well is a square ($L_x = L_y = L$)?

Solution

The energy of a particle in a 2-D well is given by Equation 3.10.24:

$$E_{n_x,n_y} = \frac{\hbar^2\pi^2}{2m}\left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2}\right)$$

and for a square well $L_x = L_y = L$ this simplifies to Equation 3.10.19:

$$E_{n_x,n_y} = \frac{\hbar^2\pi^2}{2mL^2}\left(n_x^2 + n_y^2\right)$$

When the electron absorbs an electromagnetic wave and it transitions from one state (e.g., $(n_x = 1, n_y = 2)$ or $(n_x = 2, n_y = 1)$ since are degenerate states as Table 3.10.1 demonstrates) to a higher energetic state, e.g., $(n_x = 2, n_y = 2)$.

The energy of the lowest excited state is $\frac{5h^2}{8mL^2}$ and the energy of the next higher excited state is $\frac{8h^2}{8mL^2}$, so we can equate the difference in energies of these state to the energy of the absorbed photon

$$\begin{aligned} h\nu &= E_{2,1} - E_{1,1} \\ &= \frac{5\hbar^2\pi^2}{2mL^2} - \frac{2\hbar^2\pi^2}{2mL^2} \\ &= \frac{3\hbar^2\pi^2}{2mL^2} \end{aligned}$$

Converting frequency (ν) to wavelength (λ):

$$\nu = \frac{c}{\lambda}$$

results in an expresstion relating wavelength of absorbed light to box length L:

$$\frac{hc}{\lambda} = \frac{3\hbar^2\pi^2}{2mL^2}$$

and solving for L

$$L = \sqrt{\frac{3\hbar^2\pi^2}{2mhc}\lambda}$$

or

$$\begin{aligned} L &= \sqrt{\frac{3\hbar\pi}{4mc}\lambda} \\ &= 1.91\,nm \end{aligned}$$

Exercise 3.10.2

An electron in a 2D infinite potential well needs to absorb electromagnetic wave with wavelength 4040 nm to be excited from $(n_x = 2, n_y = 2)$ state to the $(n_x = 3, n_y = 3)$ state.

a. What is the length of the box if this potential well is a square ($L_x = L_y = L$)?
b. How many wavefunction exist between these two states (do not count the start or ending states)

Contributors and Attributions

- Mark Tuckerman (New York University)
- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

This page titled 3.10: A Particle in a Two-Dimensional Box is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Mark E. Tuckerman.

- **Particle in a 2-Dimensional Box** by Mark E. Tuckerman is licensed CC BY-NC-SA 4.0.

3.11: A Particle in a Three-Dimensional Box

Learning Objectives

- To demonstrate how the particle in 1-D box problem can extend to a particle in a 3D box
- Introduction to nodal surfaces (e.g., nodal planes)

The quantum particle in the 1D box problem can be expanded to consider a particle within a higher dimensions as demonstrated elsewhere for a quantum particle in a 2D box. Here we continue the expansion into a particle trapped in a 3D box with three lengths L_x, L_y, and L_z. As with the other systems, there is NO FORCE (i.e., no potential) acting on the particles *inside* the box (Figure 3.9.1).

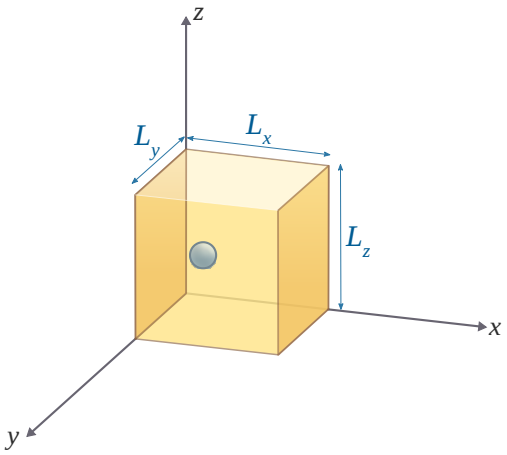


Figure 3.9.1 : A particle in 3D box with three lengths L_x, L_y, and L_z. (CC BY-NC 4.0; Ümit Kaya via LibreTexts)

The potential for the particle inside the box

$$V(\vec{r}) = 0$$

- $0 \leq x \leq L_x$
- $0 \leq y \leq L_y$
- $0 \leq z \leq L_z$
- $L_x < x < 0$
- $L_y < y < 0$
- $L_z < z < 0$

$\vec{r}$ is the vector with all three components along the three axes of the 3-D box: $\vec{r} = L_x\hat{x} + L_y\hat{y} + L_z\hat{z}$. When the potential energy is infinite, then the wavefunction equals zero. When the potential energy is zero, then the wavefunction obeys the Time-Independent Schrödinger Equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi(r) + V(r)\psi(r) = E\psi(r) \tag{3.11.1}$$

Since we are dealing with a 3-dimensional figure, we need to add the 3 different axes into the Schrondinger equation:

$$-\frac{\hbar^2}{2m}\left(\frac{d^2\psi(r)}{dx^2} + \frac{d^2\psi(r)}{dy^2} + \frac{d^2\psi(r)}{dz^2}\right) = E\psi(r) \tag{3.11.2}$$

The easiest way in solving this partial differential equation is by having the wavefunction equal to a **product** of individual function for each independent variable (e.g., the Separation of Variables technique):

$$\psi(x,y,z) = X(x)Y(y)Z(z) \tag{3.11.3}$$

Now each function has its own variable:

- $X(x)$ is a function of variable x only

- $X(x)$ is a function of variable x only
- $Y(y)$ is a function of variable y only
- $Z(z)$ is a function of variable z only

Now substitute Equation 3.11.3 into Equation 3.11.2 and divide it by the xyz product:

$$\frac{d^2\psi}{dx^2} = YZ\frac{d^2X}{dx^2} \Rightarrow \frac{1}{X}\frac{d^2X}{dx^2}$$

$$\frac{d^2\psi}{dy^2} = XZ\frac{d^2Y}{dy^2} \Rightarrow \frac{1}{Y}\frac{d^2Y}{dy^2}$$

$$\frac{d^2\psi}{dz^2} = XY\frac{d^2Z}{dz^2} \Rightarrow \frac{1}{Z}\frac{d^2Z}{dz^2}$$

$$\left(-\frac{\hbar^2}{2mX}\frac{d^2X}{dx^2}\right) + \left(-\frac{\hbar^2}{2mY}\frac{d^2Y}{dy^2}\right) + \left(-\frac{\hbar^2}{2mZ}\frac{d^2Z}{dz^2}\right) = E \tag{3.11.4}$$

E is an energy constant, and is the sum of x, y, and z. For this to work, each term must equal its own constant. For example,

$$\frac{d^2X}{dx^2} + \frac{2m}{\hbar^2}\varepsilon_x X = 0$$

Now separate each term in Equation 3.11.4 to equal zero:

$$\frac{d^2X}{dx^2} + \frac{2m}{\hbar^2}\varepsilon_x X = 0$$
$$\frac{d^2Y}{dy^2} + \frac{2m}{\hbar^2}\varepsilon_y Y = 0$$
$$\frac{d^2Z}{dz^2} + \frac{2m}{\hbar^2}\varepsilon_z Z = 0$$

Now we can add all the energies together to get the total energy:

$$\varepsilon_x + \varepsilon_y + \varepsilon_z = E \tag{3.11.5}$$

Do these equations look familiar? They should because we have now reduced the 3D box into three particle in a 1D box problems!

$$\frac{d^2X}{dx^2} + \frac{2m}{\hbar^2}E_x X = 0 \approx \frac{d^2\psi}{dx^2} = -\frac{4\pi^2}{\lambda^2}\psi \tag{3.11.6}$$

Now the equations are very similar to a 1-D box and the boundary conditions are identical, i.e.,

$$n = 1, 2, ..\infty$$

Use the normalization wavefunction equation for each variable:

$$\psi(x) = \begin{cases} \sqrt{\frac{2}{L_x}}\sin\frac{n\pi x}{L_x} & \text{if } 0 \le x \le L \\ 0 & \text{if } L < x < 0 \end{cases}$$

Normalization wavefunction equation for each variable (that substitute into Equation 3.11.3):

$$X(x) = \sqrt{\frac{2}{L_x}}\sin\left(\frac{n_x\pi x}{L_x}\right) \tag{3.11.7}$$

$$Y(y) = \sqrt{\frac{2}{L_y}}\sin\left(\frac{n_y\pi y}{L_y}\right) \tag{3.11.8}$$

$$Z(z) = \sqrt{\frac{2}{L_z}}\sin\left(\frac{n_z\pi z}{L_z}\right) \tag{3.11.9}$$

The limits of the three quantum numbers

- $n_x = 1, 2, 3, ...\infty$
- $n_y = 1, 2, 3, ...\infty$

- $n_z = 1, 2, 3, ...\infty$

For each constant use the de Broglie Energy equation:

$$\varepsilon_x = \frac{n_x^2 h^2}{8mL_x^2} \tag{3.11.10}$$

with $n_x = 1...\infty$

Do the same for variables n_y and n_z. Combine Equation 3.11.3 with Equations 3.11.7-3.11.9 to find the wavefunctions inside a 3D box.

$$\psi(r) = \sqrt{\frac{8}{V}} \sin\left(\frac{n_x \pi x}{L_x}\right) \sin\left(\frac{n_y \pi y}{L_y}\right) \sin\left(\frac{n_z \pi z}{L_z}\right) \tag{3.11.11}$$

with

$$V = \underbrace{L_x \times L_y \times L_z}_{\text{volume of box}}$$

To find the Total Energy, add Equation 3.11.10 and Equation 3.11.5.

$$E_{n_x,n_y,n_z} = \frac{h^2}{8m}\left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2}\right) \tag{3.11.12}$$

Notice the similarity between the energies a particle in a 3D box (Equation 3.11.12) and a 1D box.

Degeneracy in a 3D Cube

The energy of the particle in a 3-D cube (i.e., $L_x = L_y = L$) in the *ground state* is given by Equation 3.11.12 with $n_x = 1$, $n_y = 1$, and $n_z = 1$. This energy ($E_{1,1,1}$) is hence

$$E_{1,1,1} = \frac{3h^2}{8mL^2}$$

The ground state has only one wavefunction and no other state has this specific energy; the ground state and the energy level are said to be **non-degenerate**. However, in the 3-D cubical box potential the energy of a state depends upon the sum of the squares of the quantum numbers (Equation 3.11.11). The particle having a particular value of energy in the excited state *MAY* has several different stationary states or wavefunctions. If so, these states and energy eigenvalues are said to be **degenerate**.

For the first excited state, three combinations of the quantum numbers (n_x, n_y, n_z) are (2, 1, 1), (1, 2, 1), (1, 1, 2). The sum of squares of the quantum numbers in each combination is same (equal to 6). Each wavefunction has same energy:

$$E_{2,1,1} = E_{1,2,1} = E_{1,1,2} = \frac{6h^2}{8mL^2}$$

Corresponding to these combinations three different wavefunctions and **three** different states are possible. Hence, the first excited state is said to be three-fold or triply degenerate. The number of independent wavefunctions for the stationary states of an energy level is called as the **degree of degeneracy** of the energy level. The value of energy levels with the corresponding combinations and sum of squares of the quantum numbers

$$n^2 = n_x^2 + n_y^2 + n_z^2$$

as well as the degree of degeneracy are depicted in Table 3.9.1 .

Table 3.9.1 : Degeneracy properties of the particle in a 3-D cube with $L_x = L_y = L$.

$n_x^2+n_y^2+n_z^2$	Combinations of Degeneracy (n_x, n_y, n_z)						Total Energy (E_{n_x,n_y,n_z})	Degree of Degeneracy
3	(1,1,1)						$\frac{3h^2}{8mL^2}$	1
6	(2,1,1)	(1,2,1)	(1,1,2)				$\frac{6h^2}{8mL^2}$	3
9	(2,2,1)	(1,2,2)	(2,1,2)				$\frac{9h^2}{8mL^2}$	3
11	(3,1,1)	(1,3,1)	(1,1,3)				$\frac{11h^2}{8mL^2}$	3
12	(2,2,2)						$\frac{12h^2}{8mL^2}$	1
14	(3,2,1)	(3,1,2)	(2,3,1)	(2,1,3)	(1,3,2)	(1,2,3)	$\frac{14h^2}{8mL^2}$	6
17	(2,2,3)	(3,2,2)	(2,3,2)				$\frac{17h^2}{8mL^2}$	3
18	(1,1,4)	(1,4,1)	(4,1,1)				$\frac{18h^2}{8mL^2}$	3
19	(1,3,3)	(3,1,3)	(3,3,1)				$\frac{19h^2}{8mL^2}$	3
21	(1,2,4)	(1,4,2)	(2,1,4)	(2,4,1)	(4,1,2)	(4,2,1)	$\frac{21h^2}{8mL^2}$	6

Example 3.9.1 : Accidental Degeneracies

When is there degeneracy in a 3-D box when none of the sides are of equal length (i.e., $L_x \neq L_y \neq L_z$)?

Solution

From simple inspection of Equation 3.11.12 or Table 3.9.1 , it is clear that degeneracy originates from different combinations of n_x^2/L_x^2, n_y^2/L_y^2 and n_z^2/L_z^2 that give the same value. These will occur at common multiples of at least two of these quantities (the Least Common Multiple is one example). For example

if

$$\frac{n_x^2}{L_x^2} = \frac{n_y^2}{L_y^2}$$

there will be a degeneracy. Also degeneracies will exist if

$$\frac{n_y^2}{L_y^2} = \frac{n_z^2}{L_z^2}$$

or if

$$\frac{n_x^2}{L_x^2} = \frac{n_z^2}{L_z^2}$$

and especially if

$$\frac{n_x^2}{L_x^2} = \frac{n_y^2}{L_y^2} = \frac{n_z^2}{L_z^2}.$$

There are two general kinds of degeneracies in quantum mechanics: degeneracies due to a symmetry (i.e., $L_x = L_y$) and accidental degeneracies like those above.

Exercise 3.9.1

The 6th energy level of a particle in a 3D Cube box is 6-fold degenerate.

a. What is the energy of the 7th energy level?
b. What is the degeneracy of the 7th energy level?

References

1. Atkins, Peter. Physical Chemistry 5th Ed. USA. 1994.
2. Fitts, Donald. Principles of Quantum Mechanics. United Kingdom, Cambridge. University Press. 1999
3. McQuarrie. Donald. Physical Chemistry A Molecular Approach. Sausalito, CA. University Science Books. 1997.
4. Riggs. N. V. Quantum Chemistry. Toronto, Ontario. The Macmillan Company. 1969
5. C. A. Hollingsworth, Accidental Degeneracies of the Particle in a Box, J. Chem. Educ., 1990, 67 (12), p 999

3.11: A Particle in a Three-Dimensional Box is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.9: A Particle in a Three-Dimensional Box** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

3.E: The Schrödinger Equation and a Particle in a Box (Exercises)

Solutions to select questions can be found online.

3.2

Determine from the following operators which are linear and nonlinear:

a. $\hat{A}f(x) = f(x)^2$ [square f(x)]
b. $\hat{A}f(x) = f^*(x)$ [form the complex conjugate of f(x)]
c. $\hat{A}f(x) = 0$ [multiply f(x) by zero]
d. $\hat{A}f(x) = [f(x)]^{-1}$ [take the reciprocal of f(x)]
e. $\hat{A}f(x) = f(0)$ [evaluate f(x) at x=0]
f. $\hat{A}f(x) = \ln f(x)$ [take the log of f(x)]

3.8

Show that for a particle in a box with length a with state $n = 3$ that there are 3 locations along the x axis where the probability density is at a maximum.

3.13

What range for L is possible for σ_x given:

$$\sigma_x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$$

where L is the length of the 1-D box? Hint: Remember that σ_x is the uncertainty in the position of a particle in a box.

3.14

Using the trigonometric identity

$$\cos(2x) = 2\cos^2 x - 1$$

show that

$$\int_0^a 2\cos^2 \frac{n\pi x}{a} - 1 dx = 0$$

$$\int_0^a \cos \frac{2n\pi x}{a} dx = 0$$

$$\frac{a}{2n\pi} \sin(2n\pi) = 0$$

3.18

Is the wavefunction $\phi_n = \sqrt{\frac{2}{L}} \sin(\frac{n\pi x}{L})$ orthonormal over $0 \leq x \leq L$. Explain your reasoning.

3.22

What is the Heisenberg Uncertainty Principle? Do position and momentum follow the uncertainty principle; why or why not? If they do, what is the minimum uncertainty in the velocity of an electron if it is known to be within 1.5nm of a nucleus?

3.23

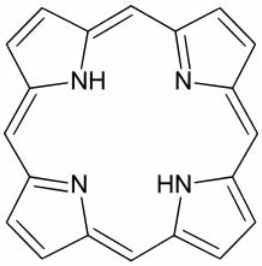
Describe the degeneracies of a two-dimensional box whose two sides have different lengths.

3.26

How many degenerate states do the first three energy levels for a three-dimensional particle in a box have if $a = b = c$?

3.27

Metal porphyrin molecules are commonly in many proteins and it has the general structure.



This molecule is planar, so we can approximate π electrons as being confined inside a square. What are energy levels and corresponding degeneracies of a particle in a square of side m? Porphyrin molecules have 18 π electrons. If the length of the molecule is 850 pm, what is the lowest energy absorption of the porphyrin molecules? (the experimental value ≈ 17,000 cm^{-1})

3.26

For a two dimensional box of width w and height $h = \sqrt{a}w$, calculate all possible energy combinations between E_{11} and E_{33} note any degeneracy.

3.32

In this problem, we will explore the quantum-mechanical problem of a free particle that is not restricted to a finite region. Remember quantized energies of a particle in a box is a direct result from the boundary conditions set by the confines of the box.

When the potential energy $V(x)$ is equal to zero and the Schrodinger equation become

$$\frac{d^2\psi}{dx^2} + \frac{2mE}{\hbar^2}\psi(x) = 0$$

The two solutions to this Schrodinger equation are

$$\psi_1(x) = A_1 e^{ikx}$$

$$\psi_2(x) = A_2 e^{-ikx}$$

where

$$k = \frac{(2mE)(1/2)}{\hbar}$$

Show that $\psi_1(x)$ and $\psi_2(x)$ are solution to the Schrodinger equation where the potential energy $V(x)$ is equal to zero

3.32

Show that E had to be a positive value, since when E is negative the wave function become unbounded for large x values

3.32

With $\hat{P}\,\psi_1(x)$ and $\hat{P}\,\psi_2(x)$ as eigenvalue equations, show that

$$\hat{P}\psi_1(x) = -i\hbar\frac{d\psi_1}{dx} = \hbar k\psi_1$$

and

$$\hat{P}\psi_2(x) = -i\hbar\frac{d\psi_2}{dx} = -\hbar k\psi_2$$

3.32

Show that $\psi_1^*\psi_1(x) = A_1^*A_1 = |A_1|^2$ and that $\psi_2^*\psi_2(x) = A_2^*A_2 = |A_2|^2$

3.33A

Assuming that a particle is characterized by a standing de Broglie wave, come up with an equation for the allowed energies of a particle in a one-dimensional box.

3.33B

Derive the lowest allowed velocity for a proton in a box of length 10^{-14} m (approximate size of nucleus), assuming the particle is described by a standing deBroglie wave.

3.33C

If a particle in a one-dimensional box is described by standing de Broglie waves within the box, derive an equation for the allowed energies. Then use that equation to find the transition energy from n=1 to n=2 given the length of the box is 350 pm and the mass of an electron is $9.109 \times 10^{-31} kg$.

3.35A

Consider the two wavefunctions

$$\psi_n(x) = \sin\frac{n\pi x}{a}$$

with even n numbers and

$$\psi_n(x) = \cos\frac{n\pi x}{a}$$

with odd n numbers.

Prove that the wavefunctions can be symmetric and antisymmetric by using the operation x to -x, a is a constant.

Given that the Schrödinger equation has the expression:

$$\hat{H}(x)\psi_n(x) = E_n\psi_n(x)$$

Through the operation x to -x, the equation now becomes:

$$\hat{H}(-x)\psi_n(-x) = E_n\psi_n(-x)$$

Show that

$$\hat{H}(x) = \hat{H}(-x)$$

is true to prove the Schrödinger equation.

3.35B

Show that the Hamiltonian for a Rigid Rotor Model is odd.

3.E: The Schrödinger Equation and a Particle in a Box (Exercises) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

4: Postulates and Principles of Quantum Mechanics

It can help us to formulate the core postulates underlying our use of quantum mechanics. If chapter 1 is the *Declaration of Independence* (from classical mechanics), then this chapter is the *Constitution* of Quantum mechanics where all our discussions rely on.

4.1: The Wavefunction Specifies the State of a System
4.2: Quantum Operators Represent Classical Variables
4.3: Observable Quantities Must Be Eigenvalues of Quantum Mechanical Operators
4.4: The Time-Dependent Schrödinger Equation
4.5: The Eigenfunctions of Operators are Orthogonal
4.6: Heisenburg Uncertainy Principle III - Commuting Operators
4.E: Postulates and Principles of Quantum Mechanics (Exercises)

4: Postulates and Principles of Quantum Mechanics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4.1: The Wavefunction Specifies the State of a System

Learning Objectives

- Introduce the first postulate of quantum mechanics
- Recognize invalid wavefunction categories

In classical mechanics, the configuration or state of a system is given by a point (x,p) in the space of coordinates and momenta. This specifies everything else in the system in a fully deterministic way, in that any observable Q that can be expressed as $Q(x,p)$ can be found, and any that cannot is irrelevant. Yet, as we have seen with the diffraction of electrons, it is **impossible** to know both the position and momentum of the electron exactly at every point along the trajectory. This is mathematically expressed as the famous position-momentum uncertainty principle. Hence, specifying a state by (x,p) in classical mechanics clearly will not work in quantum mechanics. So what specifies the state of a quantum system? This is where the **first Postulate of quantum mechanics** comes in.

Postulate I

The state of the system is **completely** specified by ψ. All possible information about the system can be found in the wavefunction ψ.

The properties of a quantum mechanical system are determined by a wavefunction $\psi(r,t)$ that depends upon the spatial coordinates of the system and time, r and t. For a single particle system, r is the set of coordinates of that particle $r=(x_1,y_1,z_1)$. For more than one particle, r is used to represent the complete set of coordinates $r=(x_1,y_1,z_1,x_2,y_2,z_2,\ldots x_n,y_n,z_n)$. Since the state of a system is defined by its properties, ψ specifies or identifies the state and sometimes is called the state function rather than the wavefunction.

What does ψ mean? This is best answered in terms of the probability density $P(x)$ that determines the probability (density) that an object in the state $\psi(x)$ will be found at position x (the Born interpretation).

$$P(x)=\psi^*(x)\psi(x) \tag{4.1.1}$$

$$=|\psi(x)|^2 \tag{4.1.2}$$

Hence, for valid (e.g., well-behaved) wavefunctions, the normalized probability in Equation 4.1.2 holds true, such that the integral over all space is equal to 1.

$$\int_{-\infty}^{\infty}\psi^*(x)\psi(x)\,dx=1 \tag{4.1.3}$$

Equation 4.1.3 means that the chance to find a particle is 100% somewhere within all space (e.g. somewhere between $-\infty$ and $+\infty$). A wavefunction is said to be **square-integrable** if Equation 4.1.3 can be satisfied (so that the Born Interpretation to be applicable).

Definition: Square-Integrable Functions

A complex-valued function, $f(x)$, is a *square-integrable* function if the integral of the square of the absolute value is finite.

$$\int_{-\infty}^{\infty}f^*(x)f(x)\,dx<\infty$$

For this to be true, the integrals of the positive and negative portions of the real and imaginary parts of $f(x)$ must both be finite.

Let us examine this set of examples in further detail in Figure 4.1.1 . The first wavefunction ψ_1 is sharply peaked at a particular value of x, and the probability density, being its square, is likewise peaked there as well. This is the wavefunction for a particle well localized at a position given by the center of the peak, as the probability density is high there, and the width of the peak is small, so the uncertainty in the position is very small.

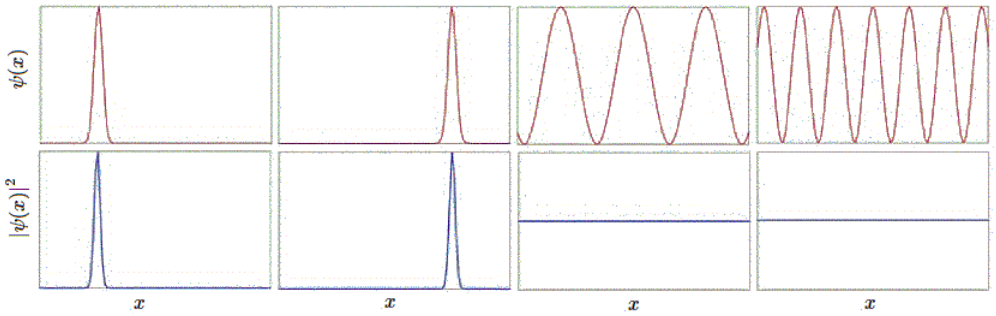


Figure 4.1.1 : Examples of good wavefunctions, $\psi(x)$, (red, top) and corresponding probability densities, $P(x)$, (blue, bottom).

The second wavefunction ψ_2 has the same peak profile, but shifted to a different position center. All of the properties of the first wavefunction hold here too, so this simply describes a particle that is well-localized at that different position. The third and fourth wavefunctions ψ_3 and ψ_4 respectively look like sinusoids of different spatial periods. The wavefunctions are actually complex of the form

$$\psi(x) = Ne^{ikx}$$

so only the real part is being plotted in Figure 4.1.1 . Note that even though the periods, k, of the oscillating wavefunctions are different,

$$\begin{aligned} P(x) &= \psi^*(x)\psi(x) \\ &= |e^{ikx}|^2 \\ &= N^2\left(e^{-ikx}\right)\left(e^{ikx}\right) \\ &= N^2 \end{aligned}$$

for all k, so the corresponding probability densities, $P(x)$, are the same except for the normalization constant (Equation 4.1.2). We saw before that it does not make a whole lot of sense to think of a sinusoidal wave as being localized in some place. Indeed, the positions for these two wavefunctions are ill-defined, so they are not well-localized, and the uncertainty in the position is large in each case. This is Heisenberg Uncertainty Principle in action.

Ill-behaved (invalid) Wavefunctions

The Born interpretation in Equation 4.1.2 means that many wavefunctions which would be acceptable mathematical solutions of the Schrödinger equation are **not acceptable** because of their implications for the physical properties of the system. To satisfy this interpretation, wavefunctions must be:

- single valued,
- continuous, and
- finite.

These aspects mean that the valid wavefunction must be one-to-one, it cannot have an undefined slope, and cannot go to $-\infty$ or $+\infty$. For example, the wavefunction must not be infinite over any finite region. If it is, then the integral in Equation 4.1.3 is equal to infinity. This implies that the particle described by such a wavefunction has a zero probability of being anywhere where the wavefunction is not infinite, but is certain to be found at all points where the wavefunction is infinite.

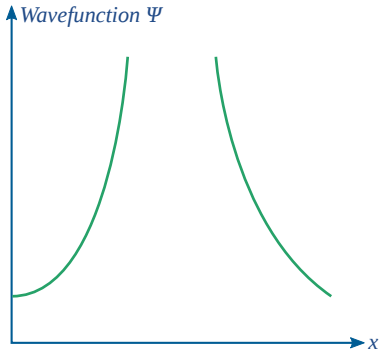


Figure 4.1.2 : This function is an invalid wavefunction since the amplitude is going to infinity over a region (the middle in this case). (CC BY-NC; Ümit Kaya via LibreTexts)

The Born interpretation also renders unacceptable solutions of the Schrödinger equation for which $|\psi(x)|^2$ has more than one value at any point. This would suggest that there were multiple different probabilities of finding the particle at that point, which is clearly absurd. The requirement that the square modulus of the wavefunction must be single-valued usually implies that the wavefunction itself must be single valued. The function in Figure 4.1.3 violates this requirement. The grey lines indicate the region where the wavefunction is multivalued.

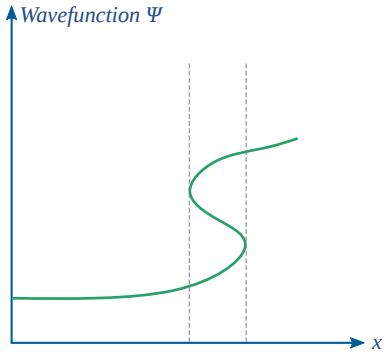


Figure 4.1.3 : This function is an invalid wavefunction since it is not a proper function (i. e., not one-to-one). (CC BY-NC; Ümit Kaya via LibreTexts)

Further restrictions arise because the wavefunction must satisfy the Schrödinger equation, which is a second-order differential equation. This implies that the second derivative of the function must exist, which implies that the first derivative of the wavefunction to exist (otherwise the second derivative is also undefined and the wavefunction cannot be a solution of the Schrödinger equation). The functions in Figure 4.1.4 are also not acceptable for these reasons.

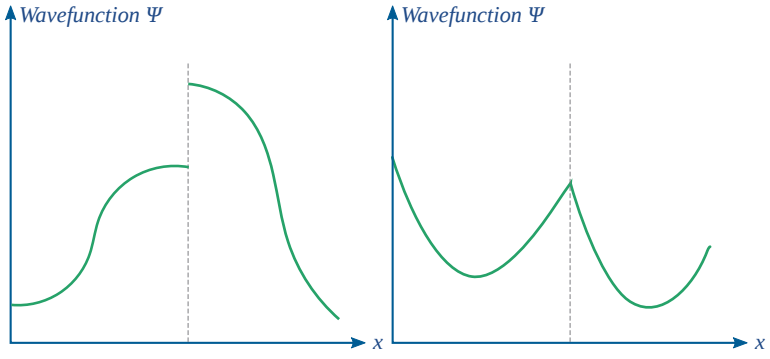


Figure 4.1.4 : These functions are invalid wavefunctions. (left) The first derivative of a function gives its gradient at a given point, and it thus exists as long as the function is continuous - only if there is a break in the function is there a point at which its first derivative does not exist. (right) The second derivative of this wavefunction is discontinuous at the point indicated, where the gradient of the line changes by more than 180°. In practice, this requirement may be somewhat flexible, particularly if the potential energy of the system shows rapid changes with distance. (CC BY-NC; Ümit Kaya via LibreTexts)

Exercise 4.1.1

Determine if each of the following functions is acceptable as a wavefunction over the indicated regions:

a. $\cos x$ over $(0, \infty)$
b. e^{x} over $(-\infty, \infty)$
c. e^{-x} over $[0, \infty)$
d. $\tan \theta$ over $[0, 2\theta]$

Contributors and Attributions

- Prof. Allan Adams, Prof. Matthew Evans, Prof. Barton Zwiebach (used with permission, MIT OpenCourseWare)

4.1: The Wavefunction Specifies the State of a System is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.1: The Wavefunction Specifies the State of a System** is licensed CC BY-NC-SA 4.0.

4.2: Quantum Operators Represent Classical Variables

Learning Objectives

- Understand how the correspondence principle argues that a unique quantum operator exist for every classical observable.
- Recognize several of the commonly used quantum operators

An observable is a dynamic variable of a system that can be experimentally measured (e.g., position, momentum and kinetic energy). In systems governed by classical mechanics, it is a real-valued function (never complex), however, in quantum physics, every observable in quantum mechanics is represented by an independent operator that is used to obtain physical information about the observable from the wavefunction. It is a general principle of quantum mechanics that there is an operator for every physical observable. For an observable that is represented in classical physics by a function $Q(x,p)$, the corresponding operator is $Q(\hat{x},\hat{p})$.

Postulate II: The Correspondence Principle

For every observable property of a system there is a corresponding quantum mechanical operator. This is often referred to as the **Correspondence Principle**.

Classical dynamical variables, such as x and p, are represented in quantum mechanics by *linear operators* which act on the wavefunction. The operator for position of a particle in three dimensions is just the set of coordinates x, y, and z, which is written as a vector, r:

$$\vec{r} = (x,y,z) \tag{4.2.1}$$

$$= x\vec{i} + y\vec{j} + z\vec{k} \tag{4.2.2}$$

The operator for a component of linear momentum is

$$\hat{p}_x = -i\hbar\frac{\partial}{\partial x} \tag{4.2.3}$$

and the operator for kinetic energy in one dimension is

$$\hat{T}_x = \left(-\frac{\hbar^2}{2m}\right)\frac{\partial^2}{\partial x^2} \tag{4.2.4}$$

and in three dimensions

$$\hat{p} = -i\hbar\nabla \tag{4.2.5}$$

and

$$\hat{T} = \left(-\frac{\hbar^2}{2m}\right)\nabla^2 \tag{4.2.6}$$

The total energy operator is called the Hamiltonian operator, $\hat{H}$ and consists of the kinetic energy operator plus the potential energy operator.

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}(x,y,z) \tag{4.2.7}$$

The Hamiltonian Operator

The Hamiltonian operator is named after the Irish mathematician William Hamilton and comes from the his formulation of Classical Mechanics that is based on the total energy:

$$\hat{H} = \hat{T} + \hat{V}$$

rather than Newton's second law,

$$\vec{F} = m\vec{a}$$

In many cases only the kinetic energy of the particles and the electrostatic or Coulomb potential energy due to their charges are considered, but in general all terms that contribute to the energy appear in the Hamiltonian. These additional terms account for such things as external electric and magnetic fields and magnetic interactions due to magnetic moments of the particles and their motion.

Table 4.2.1 : Some common Operators in Quantum Mechanics

Name	Observable Symbol	Operator Symbol	Operation
Position (in 1D)	x	$\hat{X}$	Multiply by x
Position (in 3D)	$\vec{r}$	$\hat{R}$	Multiply by $\vec{r}$
Momentum (in 1D)	p_x	$\hat{P}_x$	$-i\hbar\frac{d}{dx}$
Momentum (in 3D)	$\vec{p}$	$\hat{P}$	$-i\hbar\left[\hat{i}\,\frac{d}{dx}+\hat{j}\frac{d}{dy}+\hat{k}\frac{d}{dz}\right]$
Kinetic Energy (in 1D)	T_x	$\hat{T_x}$	$\frac{-\hbar^2}{2m}\frac{d^2}{dx^2}$
Kinetic Energy (in 3D)	T	$\hat{T}$	$\frac{-\hbar^2}{2m}\left[\frac{d^2}{dx^2}+\frac{d^2}{dy^2}+\frac{d^2}{dz^2}\right]$ Which can be simplified to $\frac{-\hbar^2}{2m}\nabla^2$
Potential Energy (in 1D)	$V(x)$	$\hat{V}(x)$	Multiply by $V(x)$
Potential Energy (in 3D)	$V(x,y,z)$	$\hat{V}(x,y,z)$	Multiply by $V(x,y,z)$
Total Energy	E	$\hat{E}$	$\frac{-\hbar^2}{2m}\nabla^2+V(x,y,z)$
Angular Momentum (x axis component)	L_x	$\hat{L_x}$	$-i\hbar\left[y\frac{d}{dz}-z\frac{d}{dy}\right]$
Angular Momentum (y axis component)	L_y	$\hat{L_y}$	$-i\hbar\left[z\frac{d}{dx}-x\frac{d}{dz}\right]$
Angular Momentum (z axis component)	L_z	$\hat{L_z}$	$-i\hbar\left[x\frac{d}{dy}-y\frac{d}{dx}\right]$

4.2: Quantum Operators Represent Classical Variables is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.2: Quantum Operators Represent Classical Variables** is licensed CC BY-NC-SA 4.0.

4.3: Observable Quantities Must Be Eigenvalues of Quantum Mechanical Operators

Learning Objectives

- To be introduced to the role of eigenvalue equations in obtaining observables from a system
- Understand how expectation values are calculated if the wavefunctions is not an eigenstate of the operator for the observable.

Recall, that we can identify the total energy operator, which is called the Hamiltonian operator, $\hat{H}$, as consisting of the kinetic energy operator plus the potential energy operator.

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}(x,y,z) \tag{4.3.1}$$

Using this notation, we write the Schrödinger Equation as

$$\hat{H}|\psi(x,y,z)\rangle = E|\psi(x,y,z)\rangle \tag{4.3.2}$$

Equation 4.3.2 says that the Hamiltonian operator operates on the wavefunction to produce the energy, which is a number, (a quantity of Joules), times the wavefunction. Such an equation, where the operator, operating on a function, produces a constant times the function, is called an **eigenvalue equation**. The function is called an **eigenfunction**, and the resulting numerical value is called the **eigenvalue**. Eigen here is the German word meaning self or own.

It is a general principle of Quantum Mechanics that there is an operator for every physical observable. A physical observable is anything that can be measured. If the wavefunction that describes a system is an eigenfunction of an operator, then the value of the associated observable is extracted from the eigenfunction by operating on the eigenfunction with the appropriate operator. The value of the observable for the system is the eigenvalue, and the system is said to be in an eigenstate. Equation 4.3.2 states this principle mathematically for the case of energy as the observable.

Postulate III: Obtaining Observables Requires Solving Eigenvalue Problems

If a system is described by the eigenfunction Ψ of an operator $\hat{A}$ then the value measured for the observable property corresponding to $\hat{A}$ will always be the eigenvalue a, which can be calculated from the eigenvalue equation.

$$\hat{A}|\Psi\rangle = a|\Psi\rangle \tag{4.3.3}$$

Consider a general real-space operator $A(x)$. When this operator acts on a general wavefunction $\psi(x)$ the result is usually a wavefunction with a completely different shape. However, there are certain special wavefunctions which are such that when A acts on them the result is just a multiple of the original wavefunction. These special wavefunctions are called *eigenstates*, and the multiples are called *eigenvalues*. Thus, if

$$A|\psi_a(x)\rangle = a|\psi_a(x)\rangle \tag{4.3.4}$$

where a is a complex number, then ψ_a is called an **eigenstate** of A corresponding to the eigenvalue a.

Suppose that A is an operator corresponding to some physical dynamical variable. Consider a particle whose wavefunction is ψ_a. The expectation of value A in this state is simply

$$\begin{aligned}\langle A\rangle &= \int_{-\infty}^{\infty} \psi_a^* A \psi_a dx \\ &= a\int_{-\infty}^{\infty} \psi_a^* \psi_a dx \\ &= a.\end{aligned}$$

where use has been made of Equation 4.3.4 and the normalization condition. Moreover,

$$\begin{aligned}\langle A^2\rangle &= \int_{-\infty}^{\infty} \psi_a^* A^2 \psi_a dx \\ &= a\int_{-\infty}^{\infty} \psi_a^* A \psi_a dx \\ &= a^2 \int_{-\infty}^{\infty} \psi_a^* \psi_a dx \\ &= a^2.\end{aligned}$$

So the variance of A is

$$\begin{aligned}\sigma_A^2 &= \langle A^2\rangle - \langle A\rangle^2 = a^2 - a^2 \\ &= 0.\end{aligned}$$

The fact that the variance is *zero* implies that every measurement of A is bound to yield the same result: namely, a. Thus, the eigenstate ψ_a is a state which is associated with a *unique* value of the dynamical variable corresponding to A. This unique value is simply the associated eigenvalue determined by Equation 4.3.4.

Expectation Values

We have seen that $|\psi(x,t)|^2$ is the probability density of a measurement of a particle's displacement yielding the value x at time t. Suppose that we made a large number of independent measurements of the displacement on an equally large number of identical quantum systems. In general, measurements made on different systems will yield different results. However, from the definition of probability, the mean of all these results is simply

$$\langle x\rangle = \int_{-\infty}^{\infty} x|\psi|^2 dx \tag{4.3.5}$$

Here, $\langle x\rangle$ is called the *expectation value* of x. Similarly the expectation value of any function of x is

$$\langle f(x)\rangle = \int_{-\infty}^{\infty} f(x)|\psi|^2 dx. \tag{4.3.6}$$

Postulate IV

The average value of an observable measurement of a state in (*normalized*) wavefunction ψ with operator $\hat{A}$ is given by the expectation value $\langle a\rangle$:

$$\langle a\rangle = \langle\psi|a|\psi\rangle \tag{4.3.7}$$

$$= \int_{-\infty}^{\infty} \psi^* \hat{A}\psi dx \tag{4.3.8}$$

If an unnormalized wavefunction where used, then Equation 4.3.8 changes to

$$\langle a\rangle = \frac{\langle\psi|a|\psi\rangle}{\langle\psi|\psi\rangle} \tag{4.3.9}$$

$$= \frac{\int_{-\infty}^{\infty} \psi^* \hat{A}\psi dx}{\int_{-\infty}^{\infty} \psi^*\psi dx} \tag{4.3.10}$$

The denominator is just the normalization requirement discussed earlier. In general, the results of the various different measurements of x will be scattered around the expectation value $\langle x\rangle$. The degree of scatter is parameterized by the quantity

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - \langle x\rangle)^2 |\psi|^2 dx \tag{4.3.11}$$

$$\equiv \langle x^2\rangle - \langle x\rangle^2, \tag{4.3.12}$$

which is known as the *variance* of x. The square-root of this quantity, σ_x, is called the *standard deviation* of x. We generally expect the results of measurements of x to lie within a few standard deviations of the expectation value (Figure 4.3.1).

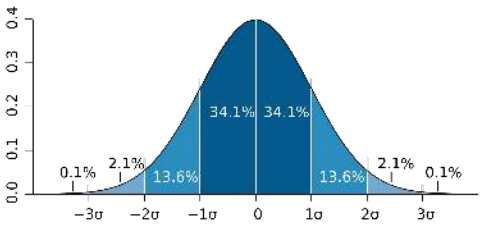


Figure 4.3.1 : A plot of a normal distribution (or bell-shaped curve) where each band has a width of 1 standard deviation. image usd with permission from Wikipedia.

Example 4.3.1

For a particle in a box in its ground state, calculate the expectation value of the

a. *position*,
b. the *linear momentum*,
c. the *kinetic energy*, and
d. the total energy

Solution

First the wavefunction needs to be defined. From the particle in the box solutions, the ground state wavefunction ($n = 1$) is

$$\psi = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right)$$

We can confirm that the wavefunction is normalized.

$$\int \psi^* \psi \, d\tau = \int_0^L \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right) \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right) dx = 1$$

Hence, the Equation 4.3.8 is the relevant equation to use.

Solution a

The expectation value of the **position** is:

$$\begin{aligned} \langle x \rangle &= \int \psi^* x \psi \, d\tau = \int_0^L \sqrt{\frac{2}{L}} x \sin\left(\frac{\pi x}{L}\right) \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right) dx \\ &= \frac{2}{L} \int_0^L x \sin^2\left(\frac{\pi x}{L}\right) dx \\ &= \frac{L}{2} \end{aligned}$$

Solution b

The expectation value of the **momentum** is:

$$\begin{aligned} \langle p \rangle &= \int \psi^* \hat{p} \psi \, d\tau = \int_0^L \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right) \left(-i\hbar \frac{d}{dx}\right) \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right) dx \\ &= \frac{2i\hbar\pi}{L^2} \int_0^L \sin\left(\frac{\pi x}{L}\right) \cos\left(\frac{\pi x}{L}\right) dx \\ &= 0 \end{aligned}$$

Solution c

The expectation value of the **kinetic energy** is:

$$\langle T \rangle = \int \psi^* \hat{K} \psi \, d\tau = \frac{2}{L} \int_0^L \sin\left(\frac{\pi x}{L}\right) \left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}\right) \sin\left(\frac{\pi x}{L}\right) dx$$
$$= \frac{\hbar^2 \pi^2}{2mL^2} \frac{2}{L} \int_0^L \sin^2\left(\frac{\pi x}{L}\right) dx$$
$$= \frac{\hbar^2 \pi^2}{2mL^2}$$

A position "on average" is in the middle of the box ($L/2$). It has equal probability of traveling towards the left or right, so the average momentum and velocity must be zero.

Solution d

The average kinetic energy must be equal to the total energy of the ground state of the particle in the box, as there is no other energy component (i.e, $V = 0$).

Expanding the Wavefunction

It is also possible to demonstrate that the eigenstates of an operator attributed to a observable form a **complete set** (*i.e.*, that any general wavefunction can be written as a linear combination of these eigenstates). However, the proof is quite difficult, and we shall not attempt it here.

In summary, given an operator $\hat{A}$, any general wavefunction, $\psi(x)$, can be written

$$\psi = \sum_i c_i \phi_i$$

where the c_i are complex weights, and the $\phi(x)$ are the properly normalized (and mutually orthogonal) eigenstates of $\hat{A}$: i.e.,

$$A\phi_i = a_i \phi_i$$

where a_i is the eigenvalue corresponding to the eigenstate ϕ_i, and

$$\int_{-\infty}^{\infty} \phi_i^* \phi_j dx = \delta_{ij}. \tag{4.3.13}$$

Here, δ_{ij} is called the **Kronecker delta-function**, and takes the value unity when its two indices are equal, and zero otherwise. It follows from Equations 4.3.10 and 4.3.13 that

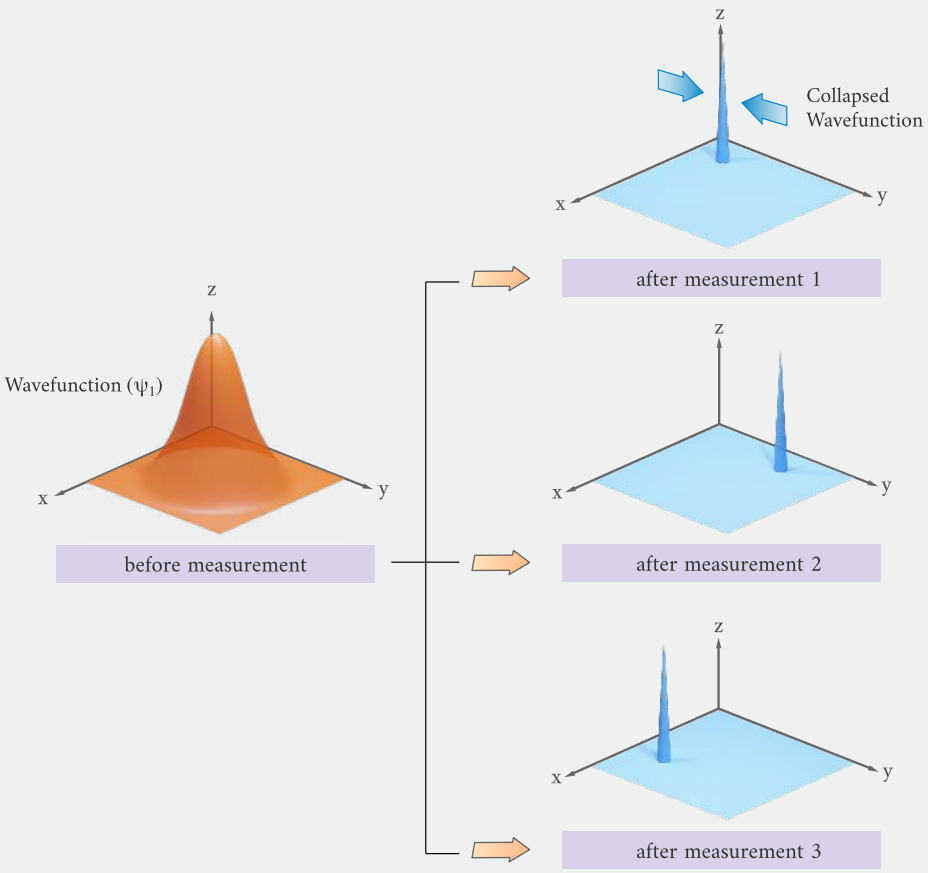
$$c_i = \int_{-\infty}^{\infty} \phi_i^* \psi dx. \tag{4.3.14}$$

Thus, the expansion coefficients in Equation 4.3.14 are easily determined, given the wavefunction ψ and the eigenstates ϕ_i. Moreover, if ψ is a properly normalized wavefunction then Equations 4.3.10 and 4.3.13 yield

$$\sum_i |c_i|^2 = 1.$$

Collapsing the Wavefunction

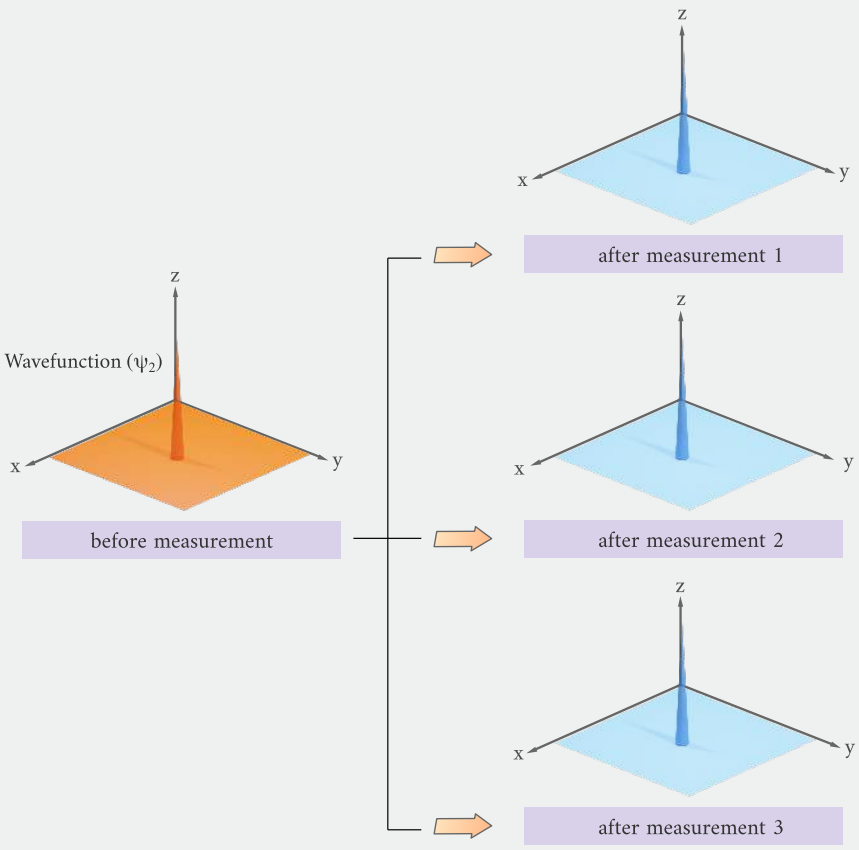
Wavefunction collapse is said to occur when a wavefunction—initially in a superposition of several eigenstates—appears to reduce to a single eigenstate (by "observation"). A particle (or a system in general) can be found in a given state $\psi(x,t)$. Suppose now a measurement is performed on the wavefuction to characterize a specific property of the system.



The measurement of position (with the position operator) effectively collapses this wavefunction that is a superposition of position eigenstates, which are delta-functions and are continuous - i.e. not quantized. Using a different measurement operation may in quantized values (CC BY-NC 4.0; Ümit Kaya via LibreTexts)

Mathematically, an operator $\hat{A}$ is associated with this measurement process, which you suppose has a *complete orthonormal set* of eigenvalues: $\{\phi_i\}$ that is typically an infinite set of functionals that depend on quantum number n. The wavefunction Ψ can be expand and a set of basis functions can be selected to specifies the wavefunction is the coefficients $\{c_n\}$ of the expansion. Therefore, if the system is perturbed, then your wavefunction will have another set of coefficients $\{c'_n\}$.

https://chem.libretexts.org/@go/page/198670



The measurement of position (with the position operator) does not collapse this wavefunction since it is in a position eigenstate in contrast to the wavefunction above. Using a different measurement operation (e.g., kinetic energy) may "collapse" the wavefunction into a different eigenstate (CC BY-NC 4.0; Ümit Kaya via LibreTexts)

If you wavefunction is in the eigenstate of the operator, then each measurement via that operator will give the same result.

Contributors and Attributions

- Richard Fitzpatrick (Professor of Physics, The University of Texas at Austin)
- Wikiversity

4.3: Observable Quantities Must Be Eigenvalues of Quantum Mechanical Operators is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.3: Observable Quantities Must Be Eigenvalues of Quantum Mechanical Operators** is licensed CC BY-NC-SA 4.0.

4.4: The Time-Dependent Schrödinger Equation

Learning Objectives

- Recognize the differences between the time-dependent and the time-independent Schrödinger equations
- To distinguish between stationary and non-stationary wavefunctions

There are two "flavors" of Schrödinger equations: the time-dependent and the time-independent versions. While the time-dependent Schrödinger equation predicts that wavefunctions can form standing waves (called stationary states), that if classified and understood, then it becomes easier to solve the time-dependent Schrödinger equation for any state. Stationary states can also be described by the time-independent Schrödinger equation (used only when the Hamiltonian is not explicitly time dependent). However, it should be noted that the solutions to the time-independent Schrödinger equation still have time dependencies.

Time-Dependent Wavefunctions

Recall that the time-independent Schrödinger equation

$$\hat{H}\psi(x) = E\psi(x) \tag{4.4.1}$$

yields the allowed energies and corresponding wavefunctions. However, it does not tell us how the system evolves in time. It would seem that something is missing, since, after all, classical mechanics tells us how the positions and velocities of a classical system evolve in time. The time dependence is given by solving Newton's second law

$$m\frac{d^2x}{dt^2} = F(x) \tag{4.4.2}$$

But where is t in quantum mechanics? First of all, what is it that must evolve in time? The answer is that the wavefunction (and associated probability density) must evolve. Suppose, therefore, that we prepare a system at $t = 0$ according to a particular probability density $p(x, 0)$ related to an amplitude $\Psi(x, 0)$ by

$$p(x, 0) = |\Psi(x, 0)|^2 \tag{4.4.3}$$

How will this initial amplitude $\Psi(x, 0)$ look at time t later? Note, by the way, that $\Psi(x, 0)$ does *not* necessarily need to be one of the eigenstates $\psi_n(x)$. To address this, we refer to the time-dependent Schrödinger equation that tells us how $\Psi(x, t)$ will evolve starting from the initial condition $\Psi(x, 0)$:

$$\hat{H}\Psi(x, t) = i\hbar\frac{\partial}{\partial t}\Psi(x, t) \tag{4.4.4}$$

It is important to know how it works physically and when it is sufficient to work with the time-independent version of the Schrödinger equation (Equation 4.4.1).

Postulate V

The time dependence of wavefunctions is governed by the Time-Dependent Schrödinger Equation (Equation 4.4.4).

Stationary States

Suppose that we are lucky enough to choose

$$\Psi(x, 0) = \psi_n(x)$$

with corresponding probability density

$$p(x, 0) = |\psi_n(x)|^2 \tag{4.4.5}$$

We will show that

$$\Psi(x, t) = \psi_n(x)e^{-iE_nt/\hbar} \tag{4.4.6}$$

From the time-dependent Schrödinger equation

$$\frac{d\Psi}{dt} = \psi_n(x)\left(\frac{-iE_n}{\hbar}\right)e^{-iE_nt/\hbar}$$
$$i\hbar\frac{d\Psi}{dt} = E_n\psi_n(x)e^{-iE_nt/\hbar}$$

Similarly

$$\begin{aligned}\hat{H}\Psi(x,t) &= e^{-iE_nt/\hbar}\hat{H}\psi_n(x)\\ &= e^{-iE_nt/\hbar}E_n\psi_n(x)\end{aligned}$$

Hence $\psi_n(x)exp(-iE_nt/\hbar)$ satisfies the Time-Dependent Schrödinger Equation (Equation 4.4.4).

Consider the probability density for this wavefunction: $p(x,t) = |\Psi(x,t)|^2$

$$\begin{aligned}p(x,t) &= \left[\psi_n(x)e^{iE_nt/\hbar}\right]\left[\psi_n(x)e^{-iE_nt/\hbar}\right]\\ &= \psi_n^2(x)e^{iE_nt/\hbar}e^{-iE_nt/\hbar}\\ &= |\psi_n(x)|^2 = p(x,0)\end{aligned}$$

the probability does not change in time and for this reason, $\psi_n(x)$ is called a *stationary state*. In such a state, the energy remains **fixed** at the well-defined value E_n.

Nonstationary States

Suppose, however, that we had chosen $\Psi(x,0)$ to be some arbitrary linear combination of the two lowest energy states:

$$\Psi(x,0) = a\psi_1(x) + b\psi_2(x) \tag{4.4.7}$$

for example

$$\Psi(x,0) = \frac{1}{\sqrt{2}}[\psi_1(x) + \psi_2(x)] \tag{4.4.8}$$

as in the previous example. Then, the probability density at time t

$$p(x,t) = |\Psi(x,t)|^2 \neq p(x,0) \tag{4.4.9}$$

For such a mixture to be possible, there must be sufficient energy in the system that there is some probability of measuring the particle to be in its excited state.

Finally, suppose we start with a state

$$\Psi(x,0) = \frac{1}{\sqrt{2}}[\psi_1(x) + \psi_2(x)]$$

and we let this state evolve in time. At any point in time, the state $\Psi(x,t)$ will be some mixture of $\psi_1(x)$ and $\psi_2(x)$, and this mixture changes with time. Now, at some specific instance in time t, we measure the energy and obtain a value E_1. What is the state of the system *just after* the measurement is made? Once we make the measurement, then we know with 100% certainty that the energy is E_1. From the above discussion, there is only one possibility for the state of the system, and that has to be the wavefunction $\psi_1(x)$, since in this state we know with 100% certainty that the energy is E_1. Hence, just after the measurement, the state *must be* $\psi_1(x)$, which means that because of the measurement, any further dependence on $\psi_2(x)$ drops out, and for all time thereafter, there is *no* dependence on $\psi_2(x)$. Consequently, any subsequent measurement of the energy would yield the value E_1 with 100% certainty. This discontinuous change in the quantum state of the system as a result of the measurement is known as the **collapse of the wavefunction**. The idea that the evolution of a system can change as a result of a measurement is one of the topics that is currently debated among quantum theorists.

The Quantum Observer Effect

The fact that measuring a quantum system changes its time evolution means that the experimenter is now coupled to the quantum system. This **observer effect** means that the act of observing will influence the phenomenon being observed. In classical mechanics, this coupling does not exist. A classical system will evolve according to Newton's laws of motion independent of whether or not we observe it. This is not true for quantum systems. The very act of observing the system changes how it evolves in time.

Put another way, by simply observing a system, we change it!

4.4: The Time-Dependent Schrödinger Equation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.4: The Time-Dependent Schrödinger Equation** by Mark E. Tuckerman is licensed CC BY-NC-SA 4.0.

4.5: The Eigenfunctions of Operators are Orthogonal

Learning Objectives

- Understand the properties of a Hermitian operator and their associated eigenstates
- Recognize that all experimental obervables are obtained by Hermitian operators

Consideration of the quantum mechanical description of the particle-in-a-box exposed two important properties of quantum mechanical systems. We saw that the eigenfunctions of the Hamiltonian operator are orthogonal, and we also saw that the position and momentum of the particle could not be determined exactly. We now examine the generality of these insights by stating and proving some fundamental theorems. These theorems use the Hermitian property of quantum mechanical operators that correspond to observables, which is discuss first.

Hermitian Operators

Since the eigenvalues of a quantum mechanical operator correspond to measurable quantities, the eigenvalues must be real, and consequently a quantum mechanical operator must be Hermitian. To prove this, we start with the premises that ψ and φ are functions, $\int d\tau$ represents integration over all coordinates, and the operator $\hat{A}$ is **Hermitian** by definition if

$$\int \psi^* \hat{A} \psi \, d\tau = \int (\hat{A}^* \psi^*) \psi \, d\tau \tag{4.5.1}$$

This equation means that the complex conjugate of $\hat{A}$ can operate on ψ^* to produce the same result after integration as $\hat{A}$ operating on φ, followed by integration. To prove that a quantum mechanical operator $\hat{A}$ is Hermitian, consider the eigenvalue equation and its complex conjugate.

$$\hat{A}\psi = a\psi \tag{4.5.2}$$

$$\hat{A}^* \psi^* = a^* \psi^* = a\psi^* \tag{4.5.3}$$

Note that $a^* = a$ because the eigenvalue is real. Multiply Equation 4.5.2 and 4.5.3 from the left by ψ^* and ψ, respectively, and integrate over the full range of all the coordinates. Note that ψ is normalized. The results are

$$\int \psi^* \hat{A} \psi \, d\tau = a \int \psi^* \psi \, d\tau = a \tag{4.5.4}$$

$$\int \psi \hat{A}^* \psi^* \, d\tau = a \int \psi \psi^* \, d\tau = a \tag{4.5.5}$$

Since both integrals equal a, they must be equivalent.

$$\int \psi^* \hat{A} \psi \, d\tau = \int \psi \hat{A}^* \psi^* \, d\tau \tag{4.5.6}$$

The operator acting on the function,

$$\hat{A}^* \int \psi^* \hat{A} \psi \, d\tau = \int \psi \hat{A}^* \psi^* \, d\tau_*$$

produces a new function. Since functions commute, Equation 4.5.6 can be rewritten as

$$\int \psi^* \hat{A} \psi d\tau = \int (\hat{A}^* \psi^*) \psi d\tau \tag{4.5.7}$$

This equality means that $\hat{A}$ is Hermitian.

Orthogonality Theorem

Eigenfunctions of a Hermitian operator are orthogonal if they have different eigenvalues.

Because of this theorem, we can identify orthogonal functions easily without having to integrate or conduct an analysis based on symmetry or other considerations.

Proof

ψ and φ are two eigenfunctions of the operator $\hat{A}$ with real eigenvalues a_1 and a_2, respectively. Since the eigenvalues are real, $a_1^* = a_1$ and $a_2^* = a_2$.

$$\hat{A}\psi = a_1\psi$$

$$\hat{A}^*\psi^* = a_2\psi^*$$

Multiply the first equation by φ^* and the second by ψ and integrate.

$$\int \psi^* \hat{A}\psi \, d\tau = a_1 \int \psi^*\psi \, d\tau$$

$$\int \psi \hat{A}^*\psi^* \, d\tau = a_2 \int \psi\psi^* \, d\tau \tag{4.5.8}$$

Subtract the two equations in Equation 4.5.8 to obtain

$$\int \psi^* \hat{A}\psi \, d\tau - \int \psi \hat{A}^*\psi^* \, d\tau = (a_1 - a_2) \int \psi^*\psi \, d\tau \tag{4.5.9}$$

The left-hand side of Equation 4.5.9 is zero because $\hat{A}$ is Hermitian yielding

$$0 = (a_1 - a_2) \int \psi^*\psi \, d\tau \tag{4.5.10}$$

If a_1 and a_2 in Equation 4.5.10 are not equal, then the integral **must** be zero. This result proves that **nondegenerate eigenfunctions** of the same operator are orthogonal.

□

Two wavefunctions, $\psi_1(x)$ and $\psi_2(x)$, are said to be *orthogonal* if

$$\int_{-\infty}^{\infty} \psi_1^* \psi_2 \, dx = 0. \tag{4.5.11}$$

Consider two eigenstates of $\hat{A}$, $\psi_a(x)$ and $\psi_{a'}(x)$, which correspond to the two *different* eigenvalues a and a', respectively. Thus,

$$A\psi_a = a\psi_a \tag{4.5.12}$$

$$A\psi_a' = a'\psi_a' \tag{4.5.13}$$

Multiplying the complex conjugate of the first equation by $\psi_{a'}(x)$, and the second equation by $\psi_{a'}^*(x)$, and then integrating over all x, we obtain

$$\int_{-\infty}^{\infty} (A\psi_a)^* \psi_{a'} dx = a \int_{-\infty}^{\infty} \psi_a^* \psi_{a'} dx, \tag{4.5.14}$$

$$\int_{-\infty}^{\infty} \psi_a^* (A\psi_{a'}) dx = a' \int_{-\infty}^{\infty} \psi_a^* \psi_{a'} dx. \tag{4.5.15}$$

However, from Equation 4.5.9, the left-hand sides of the above two equations are equal. Hence, we can write

$$(a - a') \int_{-\infty}^{\infty} \psi_a^* \psi_{a'} dx = 0.$$

By assumption, $a \neq a'$, yielding

$$\int_{-\infty}^{\infty} \psi_a^* \psi_{a'} dx = 0.$$

In other words, eigenstates of an Hermitian operator corresponding to *different* eigenvalues are automatically *orthogonal*.

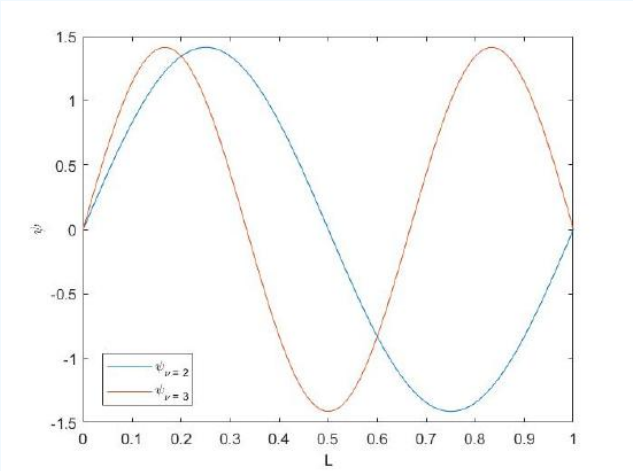
> The eigenvalues of operators associated with experimental measurements are all *real*.

Example4.5.1

Draw graphs and use them to show that the particle-in-a-box wavefunctions for $\psi(n=2)$ and $\psi(n=3)$ are orthogonal to each other.

Solution

The two PIB wavefunctions are qualitatively similar when plotted



These wavefunctions are orthogonal when

$$\int_{-\infty}^{\infty} \psi(n=2)\psi(n=3)dx = 0$$

and when the PIB wavefunctions are substituted this integral becomes

$$\int_0^L \sqrt{\frac{2}{L}} \sin\left(\frac{2n}{L}x\right) \sqrt{\frac{2}{L}} \sin\left(\frac{2n}{L}x\right) dx = ?$$

$$\frac{2}{L}\int_0^L \sin\left(\frac{2}{L}x\right) \sin\left(\frac{3}{L}x\right) = ?$$

We can expand the integrand using trigonometric identities to help solve the integral, but it is easier to take advantage of the symmetry of the integrand, specifically, the $\psi(n=2)$ wavefunction is even (blue curves in above figure) and the $\psi(n=3)$ is odd (purple curve). Their product (even times odd) is an odd function and the integral over an odd function is zero. Therefore $\psi(n=2)$ and $\psi(n=3)$ wavefunctions are orthogonal.

This can be repeated an infinite number of times to confirm the entire set of PIB wavefunctions are mutually orthogonal as the Orthogonality Theorem guarantees.

Orthogonality of Degenerate Eigenstates

Consider two eigenstates of $\hat{A}$, ψ_a and ψ'_a, which correspond to the *same* eigenvalue, a. Such eigenstates are termed **degenerate**. The above proof of the orthogonality of different eigenstates fails for degenerate eigenstates. Note, however, that *any* linear combination of ψ_a and ψ'_a is also an eigenstate of $\hat{A}$ corresponding to the eigenvalue a. Thus, even if ψ_a and ψ'_a are not orthogonal, we can always choose two linear combinations of these eigenstates which are orthogonal. For instance, if ψ_a and ψ'_a are properly normalized, we can define the **overlap integral**

$$S = \int_{-\infty}^{\infty} \psi_a^* \psi_a' dx, \tag{4.5.16}$$

It is easily demonstrated (but not here) that

$$\psi_a'' = \frac{|S|}{\sqrt{1-|S|^2}} \left(\psi_a - S^{-1}\psi_a'\right) \tag{4.5.17}$$

is a properly normalized eigenstate of $\hat{A}$, corresponding to the eigenvalue a, which is orthogonal to ψ_a. It is straightforward to generalize the above argument to three or more degenerate eigenstates. Hence, we conclude that the eigenstates of a Hermitian operator are, or can be chosen to be, **mutually orthogonal.**

Theorem: Gram-Schmidt Orthogonalization

Degenerate eigenfunctions are **not** automatically orthogonal, but can be made so mathematically via the Gram-Schmidt Orthogonalization.

The above theorem argues that if the eigenvalues of two eigenfunctions are the same then the functions are said to be degenerate and linear combinations of the degenerate functions can be formed that will be orthogonal to each other. Since the two eigenfunctions have the same eigenvalues, the linear combination also will be an eigenfunction with the same eigenvalue. The proof of this theorem shows us one way to produce orthogonal degenerate functions.

Proof

If ψ_a and ψ_a' are degenerate, but not orthogonal, we can define a new composite wavefunction $\psi_a'' = \psi_a' - S\psi_a$ where S is the overlap integral:

$$S = \langle \psi_a | \psi_a' \rangle$$

then ψ_a and ψ_a'' will be orthogonal.

$$\begin{aligned} \langle \psi_a | \psi_a'' \rangle &= \langle \psi_a | \psi_a' - S\psi_a \rangle \\ &= \cancel{\langle \psi_a | \psi_a' \rangle}^{S} - S \cancel{\langle \psi_a | \psi_a \rangle}^{1} \\ &= S - S = 0 \end{aligned}$$

$$\int \psi_a^* \psi_a d\tau$$

then ψ_a and Φ will be orthogonal.

$$\begin{aligned} \int \psi_A^* \Phi d\tau &= \int \psi_a^* (\psi_a' - S\psi_a) d\tau \\ &= \int \psi_a^* \psi_a' d\tau - S \int \psi_a^* \psi_a d\tau \\ &= S - S = 0 \end{aligned}$$

□

Exercise 4.5.2

Find N that normalizes ψ if $\psi = N(\varphi_1 - S\varphi_2)$ where φ_1 and φ_2 are normalized wavefunctions and S is their overlap integral.

$$S = \langle \varphi_1 | \varphi_2 \rangle$$

We conclude that the eigenstates of operators are, or can be chosen to be, **mutually orthogonal**.

Contributors and Attributions

- Richard Fitzpatrick (Professor of Physics, The University of Texas at Austin)

4.5: The Eigenfunctions of Operators are Orthogonal is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.5: Eigenfunctions of Operators are Orthogonal** is licensed CC BY-NC-SA 4.0.

4.6: Heisenburg Uncertainy Principle III - Commuting Operators

Learning Objectives

- To connect the Heisenberg Uncertainty principle to the commutation relations.
- Develop proficiency in calculating the commutator of two operators.

If two operators commute then both quantities can be measured at the same time with infinite precision, if not then there is a tradeoff in the accuracy in the measurement for one quantity vs. the other. This is the mathematical representation of the Heisenberg Uncertainty principle.

Commuting Operators

One important property of operators is that the order of operation matters. Thus in general

$$\hat{A}\hat{E}f(x) \neq \hat{E}\hat{A}f(x)$$

unless the two operators **commute**. Two operators commute if the following equation is true:

$$\left[\hat{A},\hat{E}\right] = \hat{A}\hat{E} - \hat{E}\hat{A} = 0$$

To determine whether two operators commute first operate $\hat{A}\hat{E}$ on a function $f(x)$. Then operate$\hat{E}\hat{A}$ the same function $f(x)$. If the same answer is obtained subtracting the two functions will equal zero and the two operators will commute.on

If two operators commute, then they can have the same set of eigenfunctions. By definition, two operators $\hat{A}$ and $\hat{B}$ commute if the effect of applying $\hat{A}$ then $\hat{B}$ is the same as applying $\hat{B}$ then $\hat{A}$, i.e.

$$\hat{A}\hat{B} = \hat{B}\hat{A}.$$

For example, the operations brushing-your-teeth and combing-your-hair commute, while the operations getting-dressed and taking-a-shower do not. This theorem is very important. If two operators commute and consequently have the same set of eigenfunctions, then the corresponding physical quantities can be evaluated or measured exactly simultaneously with no limit on the uncertainty. As mentioned previously, the eigenvalues of the operators correspond to the measured values.

If $\hat{A}$ and $\hat{B}$ commute and ψ is an eigenfunction of $\hat{A}$ with eigenvalue b, then

$$\hat{B}\hat{A}\psi = \hat{A}\hat{B}\psi = \hat{A}b\psi = b\hat{A}\psi \tag{4.6.1}$$

Equation 4.6.1 says that $\hat{A}\psi$ is an eigenfunction of $\hat{B}$ with eigenvalue b, which means that when $\hat{A}$ operates on ψ, it cannot change ψ. At most, $\hat{A}$ operating on ψ can produce a constant times ψ.

$$\hat{A}\psi = a\psi \tag{4.6.2}$$

$$\hat{B}(\hat{A}\psi) = \hat{B}(a\psi) = a\hat{B}\psi = ab\psi = b(a\psi) \tag{4.6.3}$$

Equation 4.6.3 shows that Equation 4.6.2 is consistent with Equation 4.6.1. Consequently ψ also is an eigenfunction of $\hat{A}$ with eigenvalue a.

Example 4.6.1

Do the following pairs of operators commute?

a. $\hat{A} = \dfrac{d}{dx}$ and $\hat{E} = x^2$
b. $\hat{B} = \dfrac{h}{x}$ and $\hat{C}\{f(x)\} = f(x) + 3$
c. $\hat{J} = 3x$ and $\hat{O} = x^{-1}$

General Heisenberg Uncertainty Principle

Although it will not be proven here, there is a general statement of the uncertainty principle in terms of the commutation property of operators. If two operators $\hat{A}$ and $\hat{B}$ do not commute, then the uncertainties (standard deviations σ) in the physical quantities associated with these operators must satisfy

$$\sigma_A \sigma_B \geq \left| \int \psi^* [\hat{A}\hat{B} - \hat{B}\hat{A}] \psi \, d\tau \right| \tag{4.6.4}$$

where the integral inside the square brackets is called the **commutator**, and $\left| \ \right|$ signifies the modulus or absolute value. If $\hat{A}$ and $\hat{B}$ commute, then the right-hand-side of Equation 4.6.4 is zero, so either or both σ_A and σ_B could be zero, and there is no restriction on the uncertainties in the measurements of the eigenvalues a and b. If $\hat{A}$ and $\hat{B}$ do not commute, then the right-hand-side of Equation 4.6.4 will not be zero, and neither σ_A nor σ_B can be zero unless the other is infinite. Consequently, both a and b cannot be eigenvalues of the same wavefunctions and cannot be measured simultaneously to arbitrary precision.

? Exercise4.6.1

Show that the commutator for position and momentum in one dimension equals $-i\hbar$ and that the right-hand-side of Equation 4.6.4 therefore equals $\hbar/2$ giving $\sigma_x \sigma_{px} \geq \frac{\hbar}{2}$.

Applications

Operators are very common with a variety of purposes. They are used to figure out the energy of a wavefunction using the Schrödinger Equation.

$$\hat{H}\psi = E\psi$$

They also help to explain observations made in the experimentally. An example of this is the relationship between the magnitude of the angular momentum and the components.

$$\left[\hat{L}^2, \hat{L}_x^2\right] = \left[\hat{L}^2, \hat{L}_y^2\right] = \left[\hat{L}^2, \hat{L}_z^2\right] = 0$$

However the components do not commute themselves. An additional property of commuters that commute is that both quantities can be measured simultaneously. Thus, the magnitude of the angular momentum and ONE of the components (usually z) can be known at the same time however, NOTHING is known about the other components.

> *The physical quantities corresponding to operators that commute can be measured simultaneously to any precision.*

4.6.2A ✔ Example

Determine whether the following two operators commute:

$$\hat{K} = \alpha \int [1]^{[\infty]} d[x]$$

and

$$\hat{H} = d/dx$$

4.6.2B ✔ Example

Determine whether the following two operators commute:

$$\hat{I} = 5$$

and

$$\hat{L} = \int_{[1]}^{[\infty]} d[x]$$

4.6.2C? Exercise

Show that the components of the angular momentum do not commute.

$$\hat{L}_x = -i\hbar\left[-\sin\left(\phi\frac{\delta}{\delta\theta}\right) - \cot(\Theta)\cos\left(\phi\frac{\delta}{\delta\phi}\right)\right]$$

$$\hat{L}_y = -i\hbar\left[\cos\left(\phi\frac{\delta}{\delta\theta}\right) - \cot(\Theta)\cos\left(\phi\frac{\delta}{\delta\phi}\right)\right]$$

$$\hat{L}_z = -i\hbar\frac{\delta}{\delta\theta}$$

References

1. Gohberg, I. Basic Operator Theory; Birkhäuser: Boston, 2001
2. McQuarrie, D.A. Quantum Chemistry, 2nd Edition; University Science Books:Sausalito, 2008
3. Schechter, M. Operator Methods in Quantum Mechanics; Dover Publications, 2003

4.6: Heisenburg Uncertainy Principle III - Commuting Operators is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.6: Commuting Operators Allow Infinite Precision** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0. Original source: https://web.archive.org/web/20200619182410/http://www.chemeddl.org.

4.E: Postulates and Principles of Quantum Mechanics (Exercises)

Solutions to select questions can be found online.

4.3

The function $\psi^*\psi$ has to be real, nonnegative, finite, and of definite value everywhere. Why?

4.5

Why are the following functions not acceptable wave functions for a 1D particle in a box with length a ? N is a normalization constant.

a. $\psi = N\cos\dfrac{n\pi x}{L}$

b. $\psi = \dfrac{N}{\sin\dfrac{n\pi x}{a}}$

c. $\psi = N\tan\dfrac{\pi x}{a}$

4.12

Show that the sets of functions: $\sqrt{\dfrac{2}{L}}\sin\left(\dfrac{n\pi x}{L}\right)$ where n = 1,2,3... is orthonormal.

4.13

Show that $a\cdot b\cdot c = \sum_{ik} a_i b_i c_k e_k$

$$\sum_i a_i e_i \cdot \sum_j b_j e_j \cdot \sum_k c_k e_k = \sum_{ik} a_i b_i c_k e_k$$

$$\sum_i \sum_j a_i b_j (e_i \cdot e_j) \cdot \sum_k c_k e_k = \sum_{ik} a_i b_i c_k e_k$$

$$e_i \cdot e_j = \delta_{ij} = 1$$

when $i = j$

$$\sum_i a_i b_i \cdot \sum_k c_k e_k = \sum_{ik} a_i b_i c_k e_k$$

$$\sum_{ik} a_i b_i c_k e_k = \sum_{ik} a_i b_i c_k e_k$$

4.14

Determine if the following operators commute

$$\hat{B} = \frac{d}{dx}$$

and

$$\hat{C} = x^5$$

4.15

Do the following combinations of angular momentum operators commute? Show work to the justify the answer (do not just write "yes" or no").

a. $\mathbf{L}_x$ and $\mathbf{L}_y$

b. $\mathbf{L}_y$ and $\mathbf{L}_z$

c. $\mathbf{L}_z$ and $\mathbf{L}_x$

with

$$\mathbf{L}_x = -\mathrm{i}\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$$

$$\mathbf{L}_y = -\mathrm{i}\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)$$

$$\mathbf{L}_z = -\mathrm{i}\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$$

Estimate the answer to Part C based on the pattern gathered from parts A and B; no work necessary for Part C.

4.17

For two operators to commute, what property must hold? Use the operators $\hat{L}^2$ and $\hat{L}_z$ as an example to show that this property holds.

4.21

Show that the angular momentum and kinetic energy operators commute and therefore can be measured simultaneously to arbitrary precision.

4.22

Show that the position and angular momentum operator commutes. Can the position and angular momentum be measured simultaneously to a arbitrary precision?

4.25

If both $|\Psi_n\rangle$ and $|\Psi_m\rangle$ satisfy the **time-independent** Schrödinger Equation (these are called *stationary states*)

$$|\Psi_n(x,t)\rangle = \Psi_n(x)e^{-iE_n t/\hbar}$$

and

$$|\Psi_m(x,t)\rangle = \Psi_m(x)e^{-iE_m t/\hbar}$$

show that any linear superposition of the two wavefunctions

$$|\Psi(x,t)\rangle = c_n|\Psi_n(x,t)\rangle + c_m|\Psi_m(x,t)\rangle$$

also satisfies the **time-dependent** Schrödinger Equation.

4.26

Starting with

$$\langle x \rangle = \int \psi^*(x,t) x \psi(x,t) dx$$

and the time-independent Schrödinger equation, demonstrate that

$$\frac{d\langle x \rangle}{dt} = \int \psi^* \frac{i}{\hbar} (\hat{H}x - x\hat{H}) \psi dx$$

Given that

$$\hat{H} = \frac{-\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)$$

show that

$$\hat{H}x - x\hat{H} = -2\frac{\hbar^2}{2m}\frac{d}{dx}$$
$$= -\frac{\hbar^2}{m}\frac{i}{\hbar}\hat{P}_x = -\frac{ih}{m}\hat{P}_x$$

4.28

Derive the condition on operators that arises from forcing eigenvalues to be real with complex conjugates.

4.31

Prove that the position operator is Hermitian.

4.31

Prove that the momentum operator is a Hermitian

4.32

Which of the following operators are Hermitian:

a. x,
b. d/dx
c. hd^2/dx^2
d. id^2/dx^2

4.32

Determine whether the following operators are Hermitian and whether they commute:

$$\hat{A} = i\frac{d}{dx}$$

and

$$\hat{B} = i\frac{d^2}{dx^2}$$

Given that -∞<x<∞ and the operators functions are well behaved.

4.34

Consider two wavefunctions

$$\psi_1(x) = A\sin(k_1 x) + B\cos(k_1 x) \quad (4.E.1)$$

and

$$\psi_2(x) = C\sin(k_2 x) + D\cos(k_2 x)$$

Given the boundary conditions are:

$$\psi(0) = 0$$

and

$$\frac{d\psi_1}{dx} = \frac{d\psi_2}{dx} \quad at x = 0$$

$$A + B = C, k_1(A - B) = k_2 C$$

and given a expression of

$$R = \frac{B^2}{A^2} \quad (4.E.2)$$

Derive the simplest expression of R based on the terms from the boundary conditions provided above.

4.34

A particle is moving in a field. Half-way through the field, there is a line that represents potential energy. To the left of the line, the potential energy is

$$x < 0$$

and to the right of the line the potential energy is

$$x > 0$$

. If the particle's energy is less than the potential energy line will the particle reflect when the its energy is greater than the Potential energy barrier height?

4.E: Postulates and Principles of Quantum Mechanics (Exercises) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

5: The Harmonic Oscillator and the Rigid Rotor

- **The harmonic oscillator is common:** It appears in many everyday examples: Pendulums, springs, electronics (such as the RLC circuit), standing waves on a string, etc. It's trivial to set up demonstrations of these phenomena, and we see them constantly.
- **The harmonic oscillator is intuitive:** We can picture the forces on systems such as pendulum or a plucked string. This makes it simple to study in the classroom. In contrast, there are many "everyday" examples that are not intuitive.
- **The harmonic oscillator is mathematically simple:** Math is part of physics. In studying simple harmonic motion, students can immediately use the formulas that describe its motion. These formulas are understandable: for example, the equation for frequency shows the intuitive result that increasing spring stiffness increases frequency.

5.1: A Harmonic Oscillator Obeys Hooke's Law

5.2: The Equation for a Harmonic-Oscillator Model of a Diatomic Molecule Contains the Reduced Mass of the Molecule

5.3: The Harmonic Oscillator is an Approximation

5.4: The Harmonic Oscillator Energy Levels

5.5: The Harmonic Oscillator and Infrared Spectra

5.6: The Harmonic-Oscillator Wavefunctions Involve Hermite Polynomials

5.7: Hermite Polynomials are either Even or Odd Functions

5.8: The Energy Levels of a Rigid Rotor

5.9: The Rigid Rotator is a Model for a Rotating Diatomic Molecule

5.E: The Harmonic Oscillator and the Rigid Rotor (Exercises)

5: The Harmonic Oscillator and the Rigid Rotor is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5.1: A Harmonic Oscillator Obeys Hooke's Law

The motion of two atoms in a diatomic molecule can be separated into translational, vibrational, and rotational motions. Both rotation and vibrational motions are internal motions that do not change the center of mass for the molecule (Figure 5.1.1), which is described by translational motion. Quantum translational motions can be modeled with the particle in a box model discussed previously and rotation and vibration can be modeled via the **rigid rotor** and **harmonic oscillator** models, respectively.



Figure 5.1.1 : Two atoms connected by a vibrating bond. There is a rotation around the common center of mass, and oscillation in bond distance. (CC SA-BY, 3.0; www.cleonis.nl).

Before delving into the quantum mechanical harmonic oscillator, we will introduce the classical harmonic oscillator (i.e., involving classical mechanics) to build an intuition that we will extend to the quantum world. A classical description of the vibration of a diatomic molecule is needed because the quantum mechanical description begins with replacing the classical energy with the Hamiltonian operator in the Schrödinger equation. It also is interesting to compare and contrast the classical description with the quantum mechanical picture.

The Classical Harmonic Oscillator

Simple harmonic oscillators about a potential energy minimum can be thought of as a ball rolling frictionlessly in a curved dish or a pendulum swinging frictionlessly back and forth (Figure 5.1.2). The restoring forces are precisely the same in either horizontal direction.

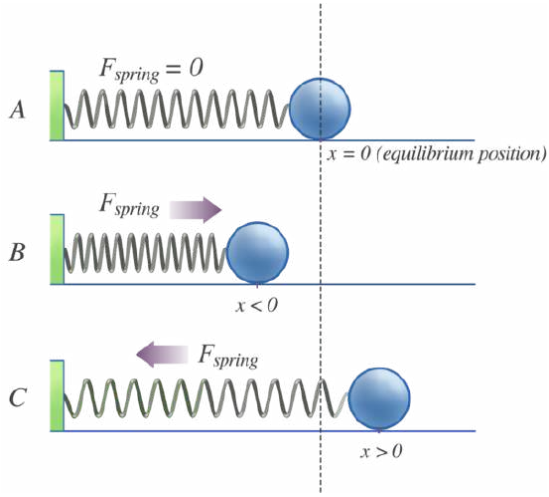


Figure 5.1.2 : A mass connected to a spring that follows Hooke's law is a system that, when displaced from its equilibrium position ($x_o = 0$), experiences a restoring force, F that is proportional to the displacement, x from the equilibrium length (x_o). (CC BY-NC; Ümit Kaya via LibreTexts)

If we consider the bond to behave like a mass on a spring (Figure 5.1.2), then this restoring force (F) is proportional to the displacement (x) from the equilibrium length (x_o) - this is *Hooke's Law*:

$$F = -kx \tag{5.1.1}$$

where k is the force constant. Hooke's Law says that the force is proportional to, but in opposite direction to, the displacement (x). The force constant reflects the **stiffness** of the spring. The idea incorporated into the application of Hooke's Law to a diatomic molecule is that when the atoms move away from their equilibrium positions, a restoring force is produced that increases proportionally with the displacement from equilibrium. The potential energy for such a system increases *quadratically* with the displacement.

$$V(x) = \frac{1}{2}kx^2 \tag{5.1.2}$$

Hooke's Law or the harmonic (i.e. quadratic) potential given by Equation 5.1.2 is an excellent *approximation* for the vibrational oscillations of molecules. The magnitude of the force constant k depends upon the nature of the chemical bond in molecular systems just as it depends on the nature of the spring in mechanical systems. The larger the force constant, the stiffer the spring or the stiffer the bond. Since it is the electron distribution between the two positively charged nuclei that holds them together, a double bond with more electrons has a larger force constant than a single bond, and the nuclei are held together more tightly.

Caution

A stiff bond with a large force constant is not necessarily a strong bond with a large dissociation energy. A harmonic oscillator has no dissociation energy since it CANNOT be broken - there is always a restoring force to keep the molecule together. This is one of many deficiencies in using the harmonic oscillator model to describe molecular vibrations.

Two atoms or one?

You may have questioned the applicability of the harmonic oscillator model involving one moving mass bound to a fix wall via a spring like in Figure 5.1.2 for the vibration of a diatomic molecule with two moving masses like in Figure 5.1.1 . It turned out the two are mathematically the same with internal vibration motion described by a single reduced particle with a reduced mass μ.

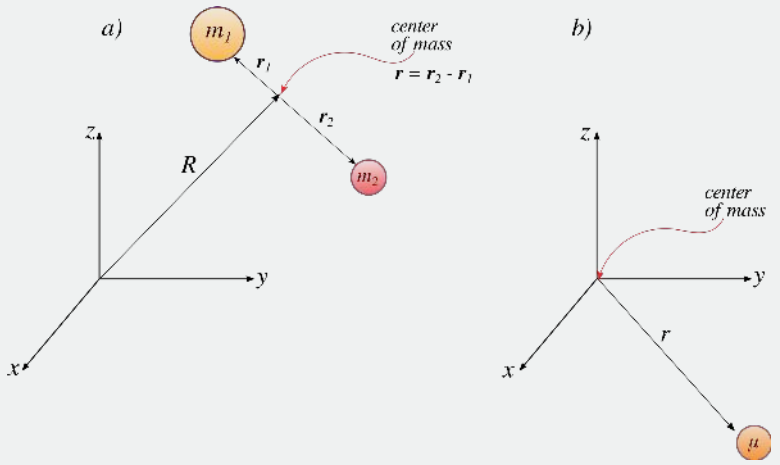


Figure 5.1.3 : Diagrams of the coordinate systems and relevant vectors for the motion of (a) a two-particle diatomic molecule with atoms of mass m_1 and m_2 and (b) the equivalent one-particle pseudo-particle of reduced mass μ. (CC BY-NC; Ümit Kaya via LibreTexts)

For a diatomic molecule, Figure 5.1.3 , the vector $\vec{r}$ corresponds to the internuclear axis. The magnitude or length of r is the bond length, and the orientation of r in space gives the orientation of the internuclear axis in space. Changes in the orientation correspond to rotation of the molecule, and changes in the length correspond to vibration. The change in the bond length from the equilibrium bond length is the **vibrational coordinate** for a diatomic molecule.

Example 5.1.1

a. Show that minus the first derivative of the harmonic potential energy function in Equation 5.1.2 with respect to x is the Hooke's Law force in Equation 5.1.1.
b. Show that the second derivative is the force constant, k.
c. At what value of x is the potential energy a minimum; at what value of x is the force zero?
d. Sketch graphs to compare the potential energy and the force for a system with a large force constant to one with a small force constant.

Solution a

Hooke's Law for a spring entails that the force applied on a spring F is equal to the force constant, $-k$ times the distance compressed or stretched, x (Equation 5.1.1). The derivative of $V(x) = 0.5kx^2$ is

$$V'(x) = (2)\left(\frac{1}{2}\right)kx = kx.$$

The negative of this is $-V'(x) = -kx$ which is exactly equal to Hooke's Law.

Solution b

The second derivative

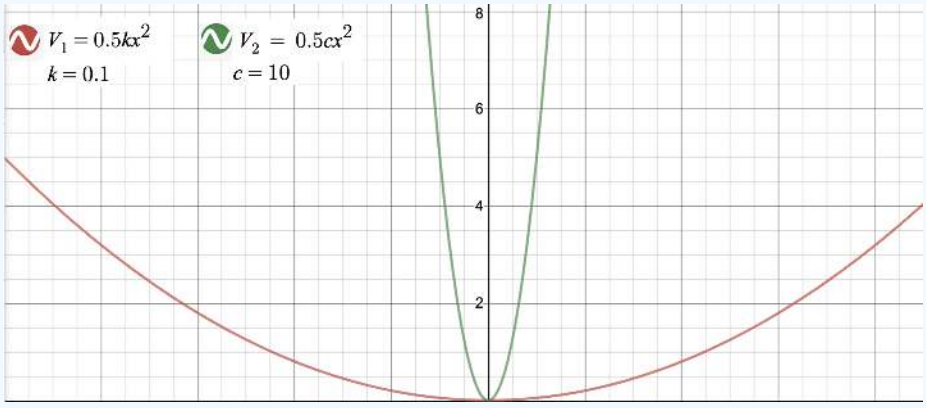
$$V''(x) = \frac{d}{dx}kx = k$$

Thus, the second derivative of this equation for potential energy is equal to the force constant, k.

Solution c

To find the minimum potential energy, it is easiest to set the first derivative equal to zero and solve for x. When $V'(x) = kx = 0$ then x must be equal to zero. Thus, the minimum potential energy is when x=0. Plugging this into Hooke's Law, $F(0) = -k(0) = 0$ so this is also the value for x when the force is zero.

Solution d



The force constant has a drastic effect on both the potential energy and the force. A system with a large force constant requires minimal change in x to have a drastic change in potential energy or force, whereas a system with a small force constant is the exact opposite phenomenon.

Solving the Harmonic Oscillator Model

The classical equation of motion for a one-dimensional simple harmonic oscillator with a particle of mass m attached to a spring having spring constant k is

$$m\frac{d^2x(t)}{dt^2} = -kx(t) \tag{5.1.3}$$

which can be rewritten in the standardform:

$$\frac{d^2x(t)}{dt^2} + \frac{k}{m}x(t) = 0 \tag{5.1.4}$$

Equation 5.1.3 is a linear second-order differential equation that can be solved by the standard method of factoring and integrating. The resulting solution to Equation 5.1.3 is

$$x(t) = x_o \sin(\omega t + \phi) \tag{5.1.5}$$

with

$$\omega = \sqrt{\frac{k}{m}} \tag{5.1.6}$$

and the momentum has time dependence

$$\begin{aligned} p &= mv \\ &= mx_o\omega\cos(\omega t + \phi) \end{aligned}$$

Figure 5.1.4 show the displacement of the bond from its equilibrium length as a function of time. Such motion is called **harmonic**.

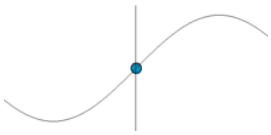


Figure 5.1.4 : Solution to the Harmonic Oscillator. Displacement (y-axis) is plotted as a function of time.

Example 5.1.2

Substitute the following functions into Equation 5.1.4 to demonstrate that they are both possible solutions to the classical equation of motion.

a. $x(t) = x_0e^{i\omega t}$
b. $x(t) = x_0e^{-i\omega t}$

where

$$\omega = \sqrt{\frac{k}{m}}$$

Note that the Greek symbol ω for frequency represents the angular frequency $2\pi\nu$.

Solution a

This requires simply placing the given function $x(t) = x_0e^{i\omega t}$ into Equation 5.1.4.

$$\begin{aligned} \frac{d^2x(t)}{dt^2} + \frac{k}{m}x(t) &= 0 \\ \frac{d^2}{dt^2}\left(x_oe^{i\omega t}\right) + \frac{k}{m}x_oe^{i\omega t} &= 0 \\ x_o\frac{d^2}{dt^2}\left(e^{i\omega t}\right) + \frac{k}{m}x_oe^{i\omega t} &= 0 \\ x_oi^2\omega^2\left(e^{i\omega t}\right) + \frac{k}{m}x_oe^{i\omega t} &= 0 \\ x_oi^2\omega^2e^{i\omega t} + \frac{k}{m}x_oe^{i\omega t} &= 0 \\ x_oi^2\omega^2 + \frac{k}{m}x_o &= 0 \\ -x_o\frac{k}{m} + \frac{k}{m}x_o = 0 \ \text{ with } \ \omega &= \sqrt{\frac{k}{m}} \end{aligned}$$

Solution b

This requires simply placing the given function $x(t) = x_0e^{-i\omega t}$ into Equation 5.1.4.

$$\frac{d^2x(t)}{dt^2} + \frac{k}{m}x(t) = 0$$

$$\frac{d^2}{dt^2}\left(x_o e^{-i\omega t}\right) + \frac{k}{m}x_o e^{-i\omega t} = 0$$

$$x_o\frac{d^2}{dt^2}\left(e^{-i\omega t}\right) + \frac{k}{m}x_o e^{-i\omega t} = 0$$

$$x_o i^2\omega^2\left(e^{-i\omega t}\right) + \frac{k}{m}x_o e^{-i\omega t} = 0$$

$$x_o i^2\omega^2 e^{-i\omega t} + \frac{k}{m}x_o e^{-i\omega t} = 0$$

$$x_o i^2\omega^2 + \frac{k}{m}x_o = 0$$

$$-x_o\frac{k}{m} + \frac{k}{m}x_o = 0 \text{ with } \omega = \sqrt{\frac{k}{m}}$$

? Exercise 5.1.1

Show that sine and cosine functions also are solutions to Equation 5.1.4.

? Exercise 5.1.2

Identify what happens to the frequency of the motion as the force constant increases in one case and as the mass increases in another case. If the force constant is increased 9-fold and the mass is increased by 4-fold, by what factor does the frequency change?

Harmonic Oscillator Energies

The energy of the vibration is the sum of the kinetic energy and the potential energy. The momentum associated with the harmonic oscillator is

$$p = m\frac{dx}{dt} \tag{5.1.7}$$

so combining Equations 5.1.7 and 5.1.2, the total energy can be written as

$$E = T + V \tag{5.1.8}$$

$$= \frac{p^2}{2m} + \frac{k}{2}x^2 \tag{5.1.9}$$

The total energy of the harmonic oscillator is equal to the maximum potential energy stored in the spring when $x = \pm A$, called the *turning points (*Figure 5.1.5). The total energy (Equation 5.1.9) is continuously being shifted between potential energy stored in the spring and kinetic energy of the mass.

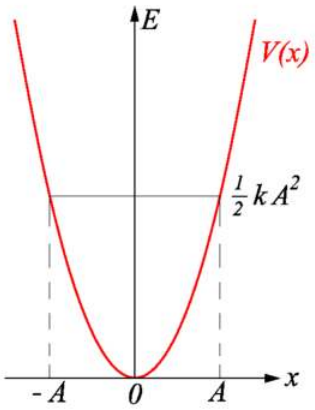


Figure 5.1.5 : The potential $V(x)$ for a harmonic oscillator. Vertical axis energy; horizontal axis displacement x. The potential energy $V(x) = ½kx^2$ is shown in red. (Paul Wormer Wikipedia).

The motion of a classical oscillator is confined to the region where its kinetic energy is nonnegative, which is what the energy relation Equation 5.1.9 says. Physically, it means that a classical oscillator can never be found beyond its turning points, and its energy depends only on how far the turning points are from its equilibrium position. The energy of a classical oscillator changes in a continuous way. The lowest energy that a classical oscillator may have is zero, which corresponds to a situation where an object is at rest at its equilibrium position. The zero-energy state of a classical oscillator simply means no oscillations and no motion at all (a classical particle sitting at the bottom of the potential well in Figure 5.1.5). When an object oscillates, no matter how big or small its energy may be, it spends the longest time near the turning points, because this is where it slows down and reverses its direction of motion. Therefore, the probability of finding a classical oscillator between the turning points is highest near the turning points and lowest at the equilibrium position. (Note that this is not a statement of preference of the object to go to lower energy. It is a statement about how quickly the object moves through various regions.)

Example5.1.3

a. What happens to the frequency of the oscillation as the vibration is excited with more and more energy?
b. What happens to the maximum amplitude of the vibration as it is excited with more and more energy?

Solution

a. Frequency

The energy of the harmonic oscillator can be written as

$$E_v = h\nu\left(v + \frac{1}{2}\right)$$

and the frequency of oscillation is $\omega = \sqrt{\frac{k}{m}}$. Notice that the frequency depends only on the stiffness (k) and reduced mass (μ) of the oscillator and not on the energy. Hence, increasing the energy of the vibrations does not affects its frequency.

b. Amplitude

The kinetic and potential terms for energy of the harmonic oscillator can be written as

$$\begin{aligned} E &= K + V \\ &= \frac{1}{2}m\omega^2A^2\sin^2\omega t + \frac{1}{2}kA^2\cos^2\omega t \end{aligned}$$

with $\omega = \sqrt{\frac{k}{m}}$ so

$$\begin{aligned} E &= \frac{1}{2}kA^2\left(\sin^2\omega t + \cos^2\omega t\right) \\ &= \frac{1}{2}kA^2 \end{aligned}$$

The maximum amplitude of the vibration will **increase** as the energy **increases.**

Exercise 5.1.3

If a molecular vibration is excited by collision with another molecule and is given a total energy E_{hit} as a result, what is the maximum amplitude of the oscillation? Is there any constraint on the magnitude of energy that can be introduced?

5.1: A Harmonic Oscillator Obeys Hooke's Law is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.1: A Harmonic Oscillator Obeys Hooke's Law** by Michael Fowler is licensed CC BY-NC-SA 4.0.

5.2: The Equation for a Harmonic-Oscillator Model of a Diatomic Molecule Contains the Reduced Mass of the Molecule

For studying the energetics of molecular vibration we take the simplest example, a diatomic heteronuclear molecule **AB**. Let the respective masses of atoms **A** and **B** be m_A and m_B. For diatomic molecules, we define the reduced mass μ_{AB} by:

$$\mu_{AB} = \frac{m_A m_B}{m_A + m_B} \tag{5.2.1}$$

Reduced mass is the representation of a two-body system as a single-body one. When the motion (displacement, vibrational, rotational) of two bodies are only under mutual interactions, the inertial mass of the moving body with respect to the body at rest can be simplified to a reduced mass.

Reduced Mass

Viewing the multi-body system as a single particle allows the separation of the motion: vibration and rotation, of the particle from the displacement of the center of mass. This approach greatly simplifies many calculations and problems.

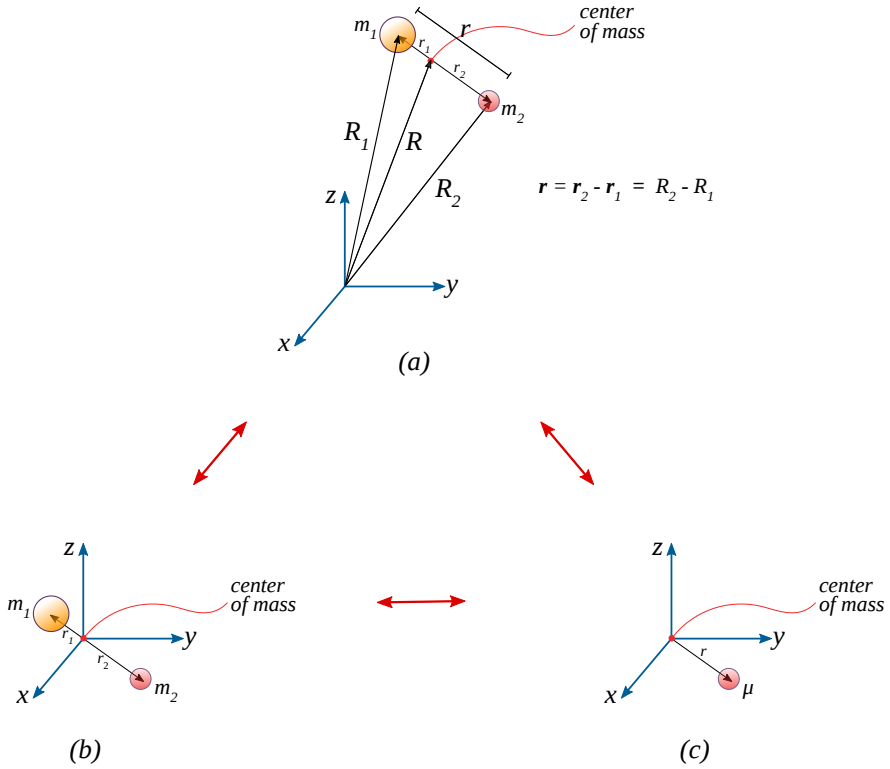


Figure 5.2.1 : a) the individual vectors to the particles m_1 and m_2 in the coordinate space and the resultant vector. b) center of mass. c) reduced mass. (CC BY-NC; Ümit Kaya via LibreTexts)

This concept is readily used in the general motion of diatomics, i.e. simple harmonic oscillator (vibrational displacement between two bodies, following Hooke's Law), the rigid rotor approximation (the moment of inertia about the center of mass of a two-body system), spectroscopy, and many other applications.

Example 5.2.1 : Reduced Mass

Determine the reduced mass of the two body system of a proton and electron with $m_{proton} = 1.6727 \times 10^{-27}\ kg$ and $m_{electron} = 9.110 \times 10^{-31}\ kg$).

The Quantum Harmonic Oscillator

The classical Harmonic Oscillator approximation is a simple yet powerful representation of the energetics of an oscillating spring system. Central to this model is the formulation of the quadratic potential energy

$$V(x) \approx \frac{1}{2}kx^2 \tag{5.2.2}$$

One problem with this classical formulation is that it is not general. We cannot use it, for example, to describe vibrations of diatomic molecules, where quantum effects are important. This require the formulation for Schrödinger Equation using Equation 5.2.2.

$$\hat{H}|\psi\rangle = \left[\frac{-\hbar^2}{2\mu}\frac{d^2}{dx^2} + \frac{1}{2}kx^2\right]|\psi\rangle = E|\psi\rangle$$

Solving this quantum harmonic oscillator is appreciably harder than solving the Schrödinger Equation for the simpler particle-in-the-box model and is outside the scope of this text. However, as with most quantum modules (and in contrast to the classical harmonic oscillator), the energies are quantized in terms of a quantum number (v in this case):

$$E_v = \hbar\left(\sqrt{\frac{k}{\mu}}\right)\left(v + \frac{1}{2}\right) \tag{5.2.3}$$

$$= h\nu\left(v + \frac{1}{2}\right) \tag{5.2.4}$$

with the natural vibrational frequency of the system given as

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \tag{5.2.5}$$

and the mass, μ, is the reduced mass of the system (Equation 5.2.1).

Warning

Be careful to distinguish ν, the symbol for the natural frequency (as a Greek nu) from v the quantum harmonic oscillator quantum number (Latin v).

Caution: Do Not Use Atomic Weights to Calculate Reduced Masses

The vibrational frequencies given by Equation 5.2.5 depend on the force constants (k) and the atomic masses of the vibrating nuclei via the reduced mass (μ). It should be clear that the substitution of one isotope of an atom in a molecule for another isotope will affect the atomic masses and therefore the reduced mass (via Equation 5.2.1) and therefore the vibrational frequencies (via Equation 5.2.5).

It is important to remember that the Periodic Table gives only atomic weights of elements, which are scaled averages of atoms normally encountered in the laboratory (Table 5.2.1). To properly discuss vibrational frequencies of molecules, we need to know (or denote) the specific isotopes in the molecule. Check Table A4 for that information.

Table 5.2.1: Atomic Mass and Isotope Composition. Consult Table A4 for more extensive table.

isotope	atomic mass (in amu)	isotopic abundance (%)
^{1}H	1.007825	99.985
^{2}H	2.0140	0.015
^{35}Cl	35.968852	75.77
^{37}Cl	36.965903	24.23
^{79}Br	78.918336	50.69
^{81}Br	80.916289	49.31

Example 5.2.1 : Isotope Effect

What are the reduced mass for $^{1}H^{35}Cl$ and $^{1}H^{37}Cl$? If the spring constants for vibrations of both molecules are equal and estimated at $478\,N/m$, what are the natural vibrational frequencies of these two molecules?

Solution

The periodic table gives an atomic weight of 35.45 amu for chlorine, but remember this is the the average of the natural abundances of different chlorine isotopes which is dictated primary by two isotopes: ^{35}Cl and ^{37}Cl. For this problem, we need the exact mass of the ^{1}H, ^{35}Cl, and ^{37}Cl isotopes. Check Table A4 for that information.

For $^{1}H^{35}Cl$:

$$\begin{aligned}\text{Reduced mass} &= \frac{m_1 m_2}{m_1 + m_2} \\ &= \frac{m_H m_{^{35}Cl}}{m_H + m_{^{35}Cl}} \\ &= \frac{(1.0078)(34.9688)}{1.0078 + 34.9688)}\,amu \\ &= 0.9796\,amu\end{aligned}$$

or when converted into kg is $1.629 \times 10^{-27}\,kg$.

For $^{1}H^{37}Cl$:

$$\begin{aligned}\text{Reduced mass} &= \frac{m_1 m_2}{m_1 + m_2} \\ &= \frac{m_H m_{^{37}Cl}}{m_H + m_{^{37}Cl}} \\ &= \frac{(1.0078)(36.965)}{1.0078 + 36.965)}\,amu \\ &= 0.9810\,amu\end{aligned}$$

or when converted into kg is $1.6291 \times 10^{-27}\,kg$. This is only 0.29% bigger.

Equation 5.2.5 is used to predict the respective vibrational frequencies of these two molecules.

For $^{1}H^{35}Cl$:

$$\begin{aligned}\nu &= \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \\ &= \frac{1}{2\pi}\sqrt{\frac{478\,N/m}{1.629 \times 10^{-27}\,kg}} \\ &= 8.6394 \times 10^{13} s^{-1}\end{aligned}$$

For $^{1}H^{37}Cl$:

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}}$$
$$= \frac{1}{2\pi}\sqrt{\frac{478\,N/m}{1.629 \times 10^{-27}\,kg}}$$
$$= 8.621 \times 10^{13} s^{-1}$$

As with the differences in the reduced masses, the differences in the vibrational frequencies of these two molecules is quite small. However, high resolution IR spectroscopy can easily distinguish the vibrations between these two molecules. Exercise 5.2.1 will demonstrate that this "**isotope effect**" is not always a small effect.

? Exercise 5.2.1 : Hydrogen Chloride

The force constant is weakly sensitive to the specific isotopes in a molecule (and we typically assume it is isotope independent). If the $k = 478\ N/m$ for both $H^{35}Cl$ and $H^{37}Cl$. What are the vibration frequencies in these two diatomic molecules.

5.2: The Equation for a Harmonic-Oscillator Model of a Diatomic Molecule Contains the Reduced Mass of the Molecule is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.2: The Equation for a Harmonic-Oscillator Model of a Diatomic Molecule Contains the Reduced Mass of the Molecule** by Eugene Lee is licensed CC BY-NC-SA 4.0.

5.3: The Harmonic Oscillator is an Approximation

Learning Objectives

- Understand the utility and limits of using the quantum harmonic oscillator as a model for molecular vibrations

The quantum harmonic oscillator is one of the most important model systems in quantum mechanics. This is due in partially to the fact that an arbitrary potential curve $V(x)$ can usually be *approximated* as a harmonic potential at the vicinity of a stable equilibrium point. Furthermore, it is one of the few quantum-mechanical systems for which an exact, analytical solution exists. Solving other potentials typically require either approximations or numerical approaches to identify the corresponding eigenstates and eigenvalues (i.e., wavefunctions and energies).

A general potential energy ($V(x)$) curve for a molecular vibration can be expanded as a Taylor series

$$V(x) = V(x_0) + \left.\frac{dV(x)}{dx}\right|_{x_0}^{x} (x - x_0) + \frac{1}{2!}\left.\frac{d^2V(x)}{dx^2}\right|_{x_0}^{x} (x - x_0)^2 + \ldots + \frac{1}{n!}\left.\frac{d^nV(x)}{dx^n}\right|_{x_0}^{x} (x - x_0)^n \tag{5.3.1}$$

It is important to note that this approximation is only good for x near x_0, and that x_0 stands for the equilibrium bond distance. $V(x)$ is often (but not always) shortened to the cubic term and can be rewritten as

$$V(x) = \frac{1}{2}kx^2 + \frac{1}{6}\gamma x^3 \tag{5.3.2}$$

where $V(x_0) = 0$, k is the harmonic force constant (harmonic term), and γ is the first anharmonic term (i.e., cubic). As Figure 5.3.2 demonstrates, the harmonic oscillator (red curve) is a good approximation for the exact potential energy of a vibration (blue curve).

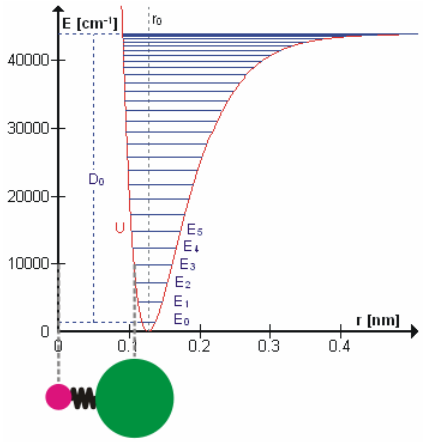


Figure 5.3.1 : The vibration of the **HCl** molecule is really an anharmonic oscillator, but can be approximated as a harmonic oscillator at low energies. In this animation, **HCl** is vibrating at the E_3 energy level. D_0 is dissociation energy here, r_0 bond length, V potential energy. Energy is expressed in wavenumbers. The hydrogen chloride molecule is attached to the coordinate system to show bond length changes on the curve. (CC BY-SA 3.0; Darekk2 via Wikipedia).

Adding anharmonic perturbations to the harmonic oscillator (Equation 5.3.2) better describes molecular vibrations. Anharmonic oscillation is defined as the deviation of a system from harmonic oscillation, or an oscillator not oscillating in simple harmonic motion. Anharmonic oscillation is described as the restoring force is no longer proportional to the displacement. Adding the cubic term (Figure 5.3.2 ; green curve) improves the harmonic oscillation approximation especially under greater displacement from equilibrium.

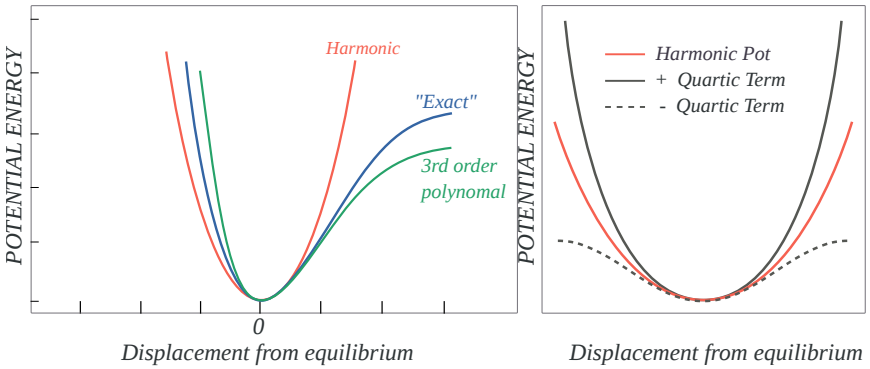


Figure 5.3.2 : (left) Deviation from simple harmonic potential approximation (red curve) of true/exact potential (blue curve) with cubic term (green). (right) Expansion with positive or negative quartic terms. (CC BY-NC; Ümit Kaya via LibreTexts)

Naturally, adding higher order anharmonic terms, like quartic terms (Figure **5.3.2;**) improve the approximation. Almost all diatomics have experimentally determined potentials for their lowest energy states. $\mathbf{H_2}$, $\mathbf{Li_2}$, $\mathbf{O_2}$, $\mathbf{N_2}$, and $\mathbf{F_2}$ with terms up to $n < 10$ determined of Equation 5.3.1.

Figure 5.3.1 shows the the general potential with (numerically) calculated energy levels (E_0, E_1 etc.). D_o is the dissociation energy, which is different from the well depth D_e. These vibrational energy levels of this plot can be calculated using the harmonic oscillator model (i.e., Equation 5.3.1 with the Schrödinger equation) and have the general form

$$E_v = \left(v + \frac{1}{2}\right) v_e - \left(v + \frac{1}{2}\right)^2 v_e x_e + \left(v + \frac{1}{2}\right)^3 v_e y_e + \textbf{higher terms}$$

where v is the vibrational quantum number and x_e and y_e are the first and second anharmonicity constants, respectively.

The $v = 0$ level is the vibrational ground state. Because this potential is less confining than a parabola used in the harmonic oscillator, the energy levels become **less widely** spaced at high excitation (Figure 5.3.1 ; top of potential).

Limitations of the Harmonic Oscillator Model for Molecular Vibrations

The harmonic oscillation is a great approximation of a molecular vibration, but has key limitations:

- Due to equal spacing of energy, all transitions occur at the **same** frequency (i.e. single line spectrum). However experimentally many lines are often observed (called overtones).
- The harmonic oscillator does not predict bond dissociation; you cannot break it no matter how much energy is introduced.

Morse Potentials are better Approximations of Vibrational Motion

A more powerful approach than just "patching up" the harmonic oscillator solution with anharmonic corrections is to adopt a different potential ($V(x)$). One such approach is the Morse potential, named after physicist Philip M. Morse, and a better approximation for the vibrational structure of the molecule than the harmonic oscillator because it explicitly includes the effects of *bond breaking* and accounts for the anharmonicity of real bonds (Figure 5.3.4).

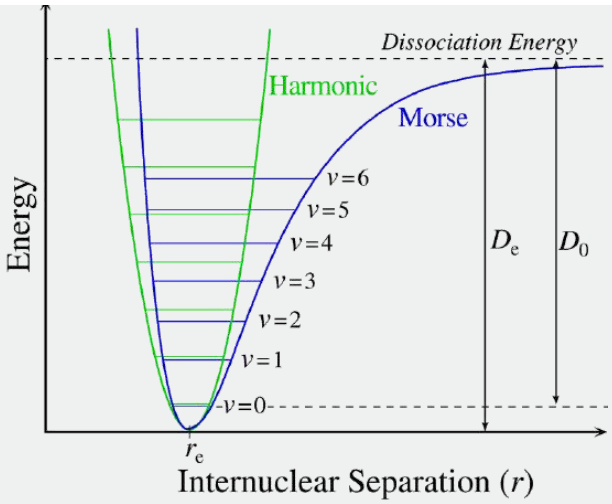


Figure 5.3.4 : The Morse potential (blue) and harmonic oscillator potential (green). Unlike the energy levels of the harmonic oscillator potential, which are evenly spaced by ħω, the Morse potential level spacing decreases as the energy approaches the dissociation energy. The dissociation energy D_e is larger than the true energy required for dissociation D_0 due to the zero point energy of the lowest ($v = 0$) vibrational level. from Wikipedia.

The Morse Potential is a good approximation to $V(x)$ and is best when looking for a general formula for all x from 0 to ∞, not just applicable for the local region around the x_o:

$$V(x) = D(1 - e^{-\beta(x-x_0)})^2 \tag{5.3.3}$$

with $V(x = x_0) = 0$ and $V(x = \infty) = D$.

The Morse Potential (Figure 5.3.4) approaches zero at infinite r_e and equals $-D_e$ at its minimum (i.e. $r = r_e$). It clearly shows that the Morse potential is the combination of a short-range repulsion term (small r values) and a long-range attractive term (large r values).

Solving the Schrödinger Equation with the Morse Potential (Equation **5.3.3**) is not trivial, but can be done analytically.

$$\hat{H}|\psi\rangle = E_n|\psi\rangle$$

with

$$\begin{aligned}\hat{H} &= \hat{T} + \hat{V} \\ &= \frac{-\hbar^2 d^2}{2m\, dx^2} + D(1 - e^{-\beta(x-x_0)})^2\end{aligned}$$

The solutions and energies for the Morse potential will not be used in this course and will not be discussed in more detail.

5.3: The Harmonic Oscillator is an Approximation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.3: The Harmonic Oscillator Approximates Molecular Vibrations** by Peter Kelly is licensed CC BY 4.0.

5.4: The Harmonic Oscillator Energy Levels

For a classical oscillator, we know exactly the position, velocity, and momentum as a function of time. The frequency of the oscillator (or normal mode) is determined by the reduced mass μ and the effective force constant k of the oscillating system and does not change unless one of these quantities is changed. There are no restrictions on the energy of the oscillator, and changes in the energy of the oscillator produce changes in the amplitude of the vibrations experienced by the oscillator.

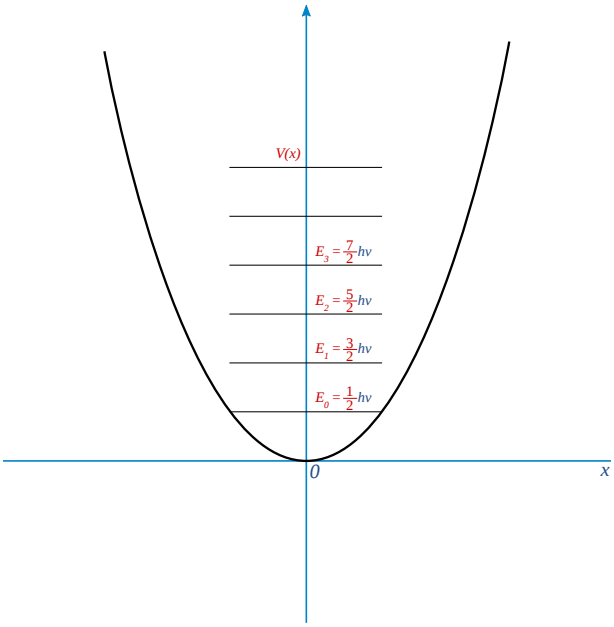


Figure 5.4.1 : Potential energy function and first few energy levels for harmonic oscillator. (CC BY-NC 4.0; Ümit Kaya via LibreTexts)

For the quantum mechanical oscillator, the oscillation frequency of a given normal mode is still controlled by the mass and the force constant (or, equivalently, by the associated potential energy function). However, the energy of the oscillator is limited to certain values. The allowed quantized energy levels are equally spaced and are related to the oscillator frequencies as given by Equation 5.4.1 and Figure 5.4.1 .

$$E_v = \left(v + \frac{1}{2}\right)\hbar\omega = \left(v + \frac{1}{2}\right)h\nu \tag{5.4.1}$$

with

$$v = 0, 1, 2, 3, \cdots \infty$$

In a quantum mechanical oscillator, we cannot specify the position of the oscillator (the exact displacement from the equilibrium position) or its velocity as a function of time; we can only talk about the probability of the oscillator being displaced from equilibrium by a certain amount. This probability is given by

$$P_{Q \to Q+dQ} = \int_{Q}^{Q+dQ} \psi_v^*(Q)\psi_v(Q)dQ \tag{5.4.2}$$

We can, however, calculate the average displacement and the mean square displacement of the atoms relative to their equilibrium positions. This average is just $\langle Q \rangle$, the expectation value for Q, and the mean square displacement is $\langle Q^2 \rangle$, the expectation value for Q^2. Similarly we can calculate the average momentum $\langle P_Q \rangle$, and the mean square momentum $\left\langle P_Q^2 \right\rangle$, but we cannot specify the momentum as a function of time.

Physically what do we expect to find for the average displacement and the average momentum? Since the potential energy function is *symmetric* around $Q = 0$, we expect values of $Q > 0$ to be equally as likely as $Q < 0$. The average value of Q therefore should be zero.

These results for the average displacement and average momentum do not mean that the harmonic oscillator is sitting still. As for the particle-in-a-box case, we can imagine the quantum mechanical harmonic oscillator as moving back and forth and therefore having an average momentum of zero. Since the lowest allowed harmonic oscillator energy, E_0, is $\frac{\hbar\omega}{2}$ and not 0, the atoms in a molecule must be moving even in the lowest vibrational energy state. This phenomenon is called the zero-point energy or the zero-point motion, and it stands in direct contrast to the classical picture of a vibrating molecule. Classically, the lowest energy available to an oscillator is zero, which means the momentum also is zero, and the oscillator is not moving.

? Exercise 5.4.1

Compare the quantum mechanical harmonic oscillator to the classical harmonic oscillator at $v = 1$ and $v = 50$.

Since the average values of the displacement and momentum are all zero and do not facilitate comparisons among the various normal modes and energy levels, we need to find other quantities that can be used for this purpose. We can use the root mean square deviation (see also root-mean-square displacement) (also known as the standard deviation of the displacement) and the root-mean-square momentum as measures of the uncertainty in the oscillator's position and momentum.

For a molecular vibration, these quantities represent the standard deviation in the bond length and the standard deviation in the momentum of the atoms from the average values of zero, so they provide us with a measure of the relative displacement and the momentum associated with each normal mode in all its allowed energy levels. These are important quantities to determine because vibrational excitation changes the size and symmetry (or shape) of molecules. Such changes affect chemical reactivity, the absorption and emission of radiation, and the dissipation of energy in radiationless transitions.

The harmonic oscillator wavefunctions form an orthonormal set; this means that all functions in the set are normalized individually

$$\int_{-\infty}^{\infty} \psi_v^*(x)\psi_v(x)dx = 1 \tag{5.4.3}$$

and are orthogonal to each other.

$$\int_{-\infty}^{\infty} \psi_{v'}^*(x)\psi_v(x)dx = 0 \text{ for } v' \neq v \tag{5.4.4}$$

The fact that a family of wavefunctions forms an orthonormal set is often helpful in simplifying complicated integrals. We will use these properties when we determine the harmonic oscillator selection rules for vibrational transitions in a molecule and calculate the absorption coefficients for the absorption of infrared radiation.

Finally, we can calculate the probability that a harmonic oscillator is in the classically forbidden region. What does this tantalizing statement mean? Classically, the maximum extension of an oscillator is obtained by equating the total energy of the oscillator to the potential energy, because at the maximum extension all the energy is in the form of potential energy. If all the energy weren't in the form of potential energy at this point, the oscillator would have kinetic energy and momentum and could continue to extend further away from its rest position. Interestingly, as we show below, the wavefunctions of the quantum mechanical oscillator extend beyond the classical limit, i.e. beyond where the particle can be according to classical mechanics.

The lowest allowed energy for the quantum mechanical oscillator is called the **zero-point energy,** $E_0 = \frac{\hbar\omega}{2}$. Using the classical picture described in the preceding paragraph, this total energy must equal the potential energy of the oscillator at its maximum extension. We define this classical limit of the amplitude of the oscillator displacement as Q_0. When we equate the zero-point energy for a particular normal mode to the potential energy of the oscillator in that normal mode, we obtain

$$\frac{\hbar\omega}{2} = \frac{kQ_0^2}{2} \tag{5.4.5}$$

> *The zero-point energy is the **lowest** possible energy that a quantum mechanical physical system may have. Hence, it is the energy of its ground state.*

Recall that k is the effective force constant of the oscillator in a particular normal mode and that the frequency of the normal mode is given by Equation 5.4.1 which is

$$\omega = \sqrt{\frac{k}{\mu}} \tag{5.4.6}$$

5.4: The Harmonic Oscillator Energy Levels is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.4: The Harmonic Oscillator Energy Levels** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

5.5: The Harmonic Oscillator and Infrared Spectra

Learning Objectives

- Understand how the quantum harmonic oscillator model can be used to interpret the infrared spectra of diatomic molecules
- Understand the origin of the transition moment integral and selection rules and how they are related

Infrared (IR) spectroscopy is one of the most common and widely used spectroscopic techniques employed mainly by inorganic and organic chemists due to its usefulness in determining structures of compounds and identifying them. Chemical compounds have different chemical properties due to the presence of different functional groups. IR spectroscopy is one of the most common and widely used spectroscopic techniques. Absorbing groups in the infrared region absorb within a certain wavelength region. The absorption peaks within this region are usually sharper when compared with absorption peaks from the ultraviolet and visible regions. In this way, IR spectroscopy can be very sensitive to determination of functional groups within a sample since different functional group absorbs different particular frequency of IR radiation. Also, each molecule has a characteristic spectrum often referred to as the fingerprint. A molecule can be identified by comparing its absorption peak to a data bank of spectra. IR spectroscopy is very useful in the identification and structure analysis of a variety of substances, including both organic and inorganic compounds. It can also be used for both qualitative and quantitative analysis of complex mixtures of similar compounds.

IR Spectroscopy

Transitions between vibrational energy levels can be induced about by absorption or emission of radiation. To understand this, knowledge of both the initial and final eigenstates is needed. The energy of the vth eigenstate of a harmonic oscillator can be written as

$$E_v = \left(v + \frac{1}{2}\right) \frac{h}{2\pi}\sqrt{\frac{k}{\mu}} \tag{5.5.1}$$

where h is Planck's constant and v is the vibrational quantum number and ranges from 0, 1, 2, 3,..., ∞. Equation 5.5.1 is often rewritten as

$$E_v = \left(v + \frac{1}{2}\right) h\nu_m \tag{5.5.2}$$

where ν_m is the vibrational frequency of the vibration. Equation 5.5.2 is often written as

$$E_v = \left(v + \frac{1}{2}\right) \hbar\omega$$

where ω is the angular frequency (i.e., $2\pi\nu$).

Transitions in vibrational energy levels can be brought about by absorption of radiation, provided the energy of the radiation ($h\nu_{photon}$) **exactly** matches the difference in energy ($\Delta E_{vv'}$) between the vibrational quantum state v to quantum state v'. This can be expressed as

$$\begin{aligned} h\nu_{photon} &= \Delta E_{vv'} \\ &= E_{v'} - E_v \\ &= \left(v' + \frac{1}{2}\right) h\nu_m - \left(v + \frac{1}{2}\right) h\nu_m \\ &= (v' - v)\, h\nu_m \end{aligned} \tag{5.5.3}$$

Let's consider only transitions between adjacent eigenstates (discussed in more details below) so

$$v' - v = \pm 1$$

which is positive if an IR photon is absorbed and negative if it is emitted. For the absorption of a IR photon, Equation 5.5.3 simplifies to

$$h\nu_{photon} = h\nu_m$$
$$= \frac{h}{2\pi}\sqrt{\frac{k}{\mu}} \tag{5.5.4}$$

The frequency of radiation ν_{photon} that will bring about this change is **identical** to the classical vibrational frequency of the bond ν.

The Wavenumber as a Unit of Energy

The cm^{-1} is the wavenumber scale and it can also be defined as 1/wavelength in cm. A wavenumber is often used due to its direct relationship with both frequency and energy. The frequency of the absorbed radiation causes the molecular vibrational frequency for the absorption process:

$$\bar{\nu}(cm^{-1}) = \frac{1}{\lambda(\mu m)} \times 10^4 \left(\frac{\mu m}{cm}\right) \tag{5.5.5}$$
$$= \frac{\nu\,(Hz)}{c\,(cm/s)} \tag{5.5.6}$$

Equation 5.5.3 can be modified so that the radiation can be expressed in wavenumbers

$$\tilde{\nu}_m = \frac{1}{2\pi c}\sqrt{\frac{k}{\mu}} \tag{5.5.7}$$

where c is the velocity of light (cm/s) and $\tilde{\nu}$ is the wavenumber of an absorption maximum (cm^{-1}).

Example 5.5.1: Hydrogen Halides

The force constants for typical diatomic molecules are in the range between 400 to 2000 $N \cdot m^{-1}$.

Molecule	HF	HCl	HBr	HI	CO	NO
Force constant, k ($N.m^{-1}$)	970	480	410	320	1860	1530

For the diatomic molecules listed above, calculate the following:

a. angular frequency ($rad \cdot s^{-1}$)
b. natural frequency (Hz)
c. period (s)
d. separation between energy levels
e. wavelength λ of the electromagnetic radiation absorbed in the transition $v = 0 \rightarrow v = 1$.

Solution

For **HCl**:

a. angular frequency

$$\omega = 5.45 \times 10^{14} rad \cdot s^{-1}$$

b. natural frequency:

$$\nu = \frac{\omega}{2\pi} = 8.68 \times 10^{13}$$

c. period (s):

$$T = \frac{1}{\nu} = 1.15 \times 10^{-14}$$

d. separation between energy levels:

$$\Delta E = E_{v=1} - E_{v=0} = \hbar\omega = 5.75 \times 10^{-20}\ J$$

e. wavelength λ of the electromagnetic radiation absorbed in the transition $v = 0 \rightarrow v = 1$:

$$3.46 \times 10^{-6}\, m$$

The electromagnetic radiation released (and absorbed) for vibrations is primarily in the infrared (IR) part of the spectrum. Calculating the above properties for the other molecules remains as an exercise.

Note: We did not designate the specific isotopes in the hydrogen halides above because that information was not required to solve these questions.

IR deals with the interaction between a molecule and radiation from the electromagnetic region ranging (4000- 40 cm^{-1}). The IR region of the electromagnetic spectrum ranges in wavelength from 2 -15 μm. Conventionally the IR region is subdivided into three regions (Table 5.5.1): near IR, mid IR and far IR. Most of the IR used originates from the mid IR region.

Table 5.5.1: Regions of the IR Spectrum

Region	Wavelength	Wavenumbers ($\tilde{\nu}$), cm^{-1}	Frequencies (v), HZ
Near	0.78 -2.5	12800 - 4000	3.8 x 10^{14} - 1.2 x 10^{14}
Middle	2.5 - 50	4000 - 200	3.8 x 10^{14} - 1.2 x 10^{14}
Far	50 -100	200 -10	3.8 x 10^{14} - 1.2 x 10^{14}
Most Used	2.5 -15	4000 -670	3.8 x 10^{14} - 1.2 x 10^{14}

? Example 5.5.2: Hydrogen Chloride

The $H^{35}Cl$ equilibrium bond length is 0.127 nm and the $v = 0$ to $v = 1$ transition is observed in the infrared at 2,886 cm^{-1}. Compute the vibrational energy of $H^{35}Cl$ in its lowest state.

a. Compute the classical limit for the stretching of the HCl bond from its equilibrium length in this state.
b. What percent of the equilibrium bond length is this extension?

Solution

The H—Cl bond length is $0.127\,nm = 1.27 \times 10^{-10} m$ and the IR transition is observed at $2886\,cm^{-1}$. We first convert this to Hertz

$$2886\,cm^{-1}(3 \times 10^{10} cm/s) = 8.646 \times 10^{13} s^{-1} = 8.646 \times 10^{13}\,Hz$$

The equations for the energy of the vth eigenstate of a harmonic oscillator is

$$E_v = (v + 1/2)\hbar\omega$$

So the ground state ($v = 0$) energy is

$$E_0 = \frac{1}{2}\hbar\omega$$

with

$$\begin{aligned}\omega &= 2\pi\nu \\ &= 5.54 \times 10^{14}\, rad/s\end{aligned}$$

We can then extra the ground state vibrational energy (aka zero-point energy) from

$$\begin{aligned}E_o &= \frac{1}{2}\hbar\omega \\ &= \frac{1}{2}\hbar(5.54 \times 10^{14} rad/s) \\ &= 2.916 \times 10^{-20}\ J\end{aligned}$$

The classical limit of the stretch is denoted as Q_0, this can be equated as potential energy in relation to the total E_0 found above as, at E_0, all of the energy would be potential energy in the form of the stretch. In comparison to the classic spring potential

$$V = \frac{1}{2}kQ_0^2$$

As described above, we can relate the two as

$$\frac{1}{2}\hbar\omega = \frac{1}{2}kQ_0^2$$

or,

$$2.916 \times 10^{-20} J = \frac{1}{2}481Q_0^2$$

where $Q_0 = 1.10 \times 10^{-11} m = 0.0110\, nm$.

Lastly, this classical limit to length can be compared to the equilibrium bond length by a simple relation of

$$\begin{aligned} \text{Percent bond length} &= \frac{Q_0}{x_{eq}} \times 100\% \\ &= \frac{0.0110\, nm}{0.127\, nm} \times 100\% \\ &= 8.66\% \end{aligned}$$

Selection Rules for IR Transitions

Photons can be absorbed or emitted, and the harmonic oscillator can transition from one vibrational energy state to another. Which transitions between vibrational states are allowed? When discussing the Bohr hydrogen atom in Chapter 1, we identified multiple transitions between different states as being responsible for the line spectra emission of the hydrogen atom. No limitations were identified so that any transition between any starting state and ending state was possible and this resulted in a complex pattern in the spectra. As we will identify later on, when considering mulit-electron atoms, there are **selection rules** that limits which transitions will be observed (i.e., are allowed).

A similar situation exist for the harmonic oscillator and IR transitions and a set of selection rules must be satisfied to ensure a particular transition is allowed. Selection rules result from evaluating a **transition moment integral** that expresses the probability of a transition from the v to the v' eigenstates:

$$\mu_T = \langle \psi_{v'} | \hat{\mu}(x) | \psi_v \rangle \tag{5.5.8}$$

$$= \int_{-\infty}^{\infty} \psi_{v'}^*(x) \hat{\mu}(x) \psi_v(x) dx \tag{5.5.9}$$

When the transition moment integral is zero for two specific eigenstates (i.e., $\psi_v(x)$ and $\psi_v'(x)$), then the probability of observing that transition is zero. In this case, this transition is called a **forbidden transition** (i.e., it will not be obversed). Transitions with low probabilities are call "weakly allowed transitions" since they will have small amplitudes in spectra and transitions with high probabilities are "strongly allowed transitions" and will exhibit high amplitudes in spectra.

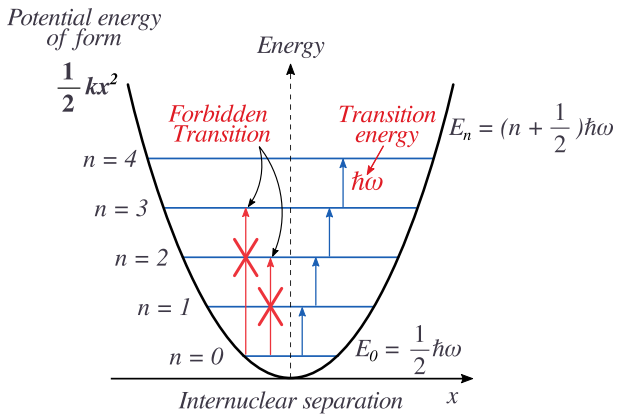


Figure 5.5.1: Possible transitions for a harmonic oscillator. Allowed transitions are labeled with blue arrows and forbidden transitions are labeled with red arrows. (CC BY-NC; Ümit Kaya via LibreTexts)

To evaluate Equation 5.5.9, we need to express the dipole moment operator, $\hat{\mu}$, in terms of the magnitude of the vibration x. The dipole moment operator is defined as

$$\hat{\mu} = \sum_{electrons} e\vec{r} + \sum_{nuclei} q\vec{R} \tag{5.5.10}$$

where the two sums are over all the electrons and nuclei and involve the particle charge ($-e$ or q) multiplying the position vector ($\vec{r}$ or $\vec{R}$, respectively). We can obtain this dipole moment operator in terms of the magnitude of the displacement coordinate, x, in a simple way by using a **Taylor series expansion** for the dipole moment around the equlibrium position ($x = 0$).

$$\hat{\mu}(x) = \mu_{x=0} + \left.\frac{d\mu(x)}{dx}\right|_{x=0} x + \left.\frac{d^2\mu(x)}{dx^2}\right|_{x=0} x^2 + \cdots \tag{5.5.11}$$

Retaining only the first two terms and substituting into Equation 5.5.9 produces

$$\mu_T = \mu_{x=0} \int_{-\infty}^{\infty} \psi_{v'}(x)\psi_v(x)dx + \left.\frac{d\mu(x)}{dx}\right|_{x=0} \int_{-\infty}^{\infty} x\psi_{v'}^*(x)\psi_v(x)dx \tag{5.5.12}$$

where $\mu_{x=0}$ is the **dipole moment** of the molecule when the nuclei are at their equilibrium positions, and

$$\left.\frac{d\mu(x)}{dx}\right|_{x=0}$$

is the *linear* change in the dipole moment due to the displacement of the nuclei in the normal mode. The derivative is the linear change because it multiplies x and not a higher power of x in Equation 5.5.11. Both μ and

$$\left.\frac{d\mu(x)}{dx}\right|_{x=0}$$

are moved outside of the integral because they are constants that no longer depend on x because they are evaluated at $x = 0$.

The integral in the first term in Equation 5.5.12 is 0 because any two harmonic oscillator wavefunctions are *orthogonal*. The integral in the second term of Equation 5.5.12 is zero except when $v' = v \pm 1$ as demonstrated in Exercise 5.5.1 . Also note that the second term is zero if

$$\left.\frac{d\mu(x)}{dx}\right|_{x=0} = 0 \tag{5.5.13}$$

Hence we can identify two "rules" that need to be satisfied for an IR photon to be absorb (or emitted) by a vibrating molecule.

Rules for IR Absorption

For IR absorption to occur two conditions must be met:

1. There must be a **change** in the dipole moment of the molecule as a result of a molecular vibration (or rotation). The change (or oscillation) in the dipole moment allows interaction with the alternating electrical component of the IR radiation. Symmetric molecules (or bonds) do not absorb IR radiation since there is no dipole moment.
2. If the frequency of the radiation ν_{photon} matches the natural frequency of the vibration (ν_m), the IR photon can be absorbed and the amplitude of the vibration increases.

? Exercise 5.5.1

Use one of the Hermite polynomial recursion relations to verify that the second integral in Equation 5.5.12 is 0 unless $v' = v \pm 1$.

If we are to observe absorption of infrared radiation due to a vibrational transition in a molecule, the transition moment cannot be zero. This condition requires that the dipole moment derivative Equation 5.5.13 **cannot** be zero and that the vibrational quantum number change by one unit. The normal coordinate motion must cause the dipole moment of the molecule to change in order for a molecule to absorb infrared radiation. If the normal coordinate oscillation does not cause the dipole moment to change then $\mu_T = 0$ and no infrared absorption is observed. So we can

$$\underbrace{\Delta v = \pm 1}_{\text{For allowed transitions}} \tag{5.5.14}$$

Consider oxygen and nitrogen molecules. Because they are symmetrical, their dipole moments are zero, $\mu = 0$. Since the vibrational motion (only bond stretching for a diatomic molecule) preserves this symmetry, the change in the dipole moment due to the vibrational motion also is zero, $\frac{d\mu(x)}{dx} = 0$. Consequently, oxygen and nitrogen do not absorb infrared radiation as a result of vibrational motion.

? Exercise 5.5.2

Explain why the molar absorptivity or molar extinction coefficients (ϵ) in Beer's Law for the IR absorption of some vibrations are greater than others.

The case $v' = v + 1$ corresponds to going from one vibrational state to a higher energy one by absorbing a photon with energy $h\nu$. The case $v' = v - 1$ corresponds to a transition that emits a photon with energy $h\nu$. In the harmonic oscillator model infrared spectra are very simple; only the fundamental transitions, $\Delta = \pm 1$, are allowed. The associated transition energy is $\hbar\omega$, according to Equation 5.5.4. The transition energy is the change in energy of the oscillator as it moves from one vibrational state to another, and it equals the photon energy.

$$\begin{aligned} \Delta E &= E_{final} - E_{initial} \\ &= h\upsilon_{photon} \\ &= \hbar\omega_{oscillator} \end{aligned}$$

For perfect harmonic oscillators, the only possible allowed transitions are $\Delta = \pm 1$ with all other transitions forbidden (Figure 5.5.1). This conclusion predicts that the vibrational absorption spectrum of a diatomic molecule consists of a single line since the energy levels are equally spaced in the harmonic oscillator model (Figure 5.5.1). If the vibration were anharmonic, then the levels would not be equally spaced and then transitions from $v = 0$ to $v = 1$ and from $v = 1$ to $v = 2$, etc. would occur at different frequencies.

> *Only the fundamental transitions, $\Delta = \pm 1$, are observed in infrared spectra within harmonic oscillator model.*

The actual IR spectrum is more complex, especially at high resolution. There is a fine structure due to the rotational states of the molecule. These states will be discussed in the next chapter. The spectrum is enriched further by the appearance of lines due to transitions corresponding to $\Delta = \pm n$ where $n > 1$. These transitions are called **overtone** transitions and their appearance in spectra despite being forbidden in the harmonic oscillator model is due to the anharmonicity of molecular vibrations. Anharmonicity means

the potential energy function is not strictly the harmonic potential. The first overtone, $\Delta v = 2$, generally appears at a frequency slightly less than twice that of the fundamental, i.e. the frequency due to the $\Delta v = 1$ transition.

? Exercise 5.5.3: Hydrogen Chloride

Compute the approximate transition frequencies in wavenumber units for the first and second overtone transitions in $\mathbf{HCl}$ given that the fundamental is at 2,886 cm^{-1}.

5.5: The Harmonic Oscillator and Infrared Spectra is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.5: The Harmonic Oscillator and Infrared Spectra** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

5.6: The Harmonic-Oscillator Wavefunctions Involve Hermite Polynomials

Learning Objectives

- The Equation for a Harmonic-Oscillator Model of a Diatomic Molecule Contains the Reduced Mass of the Molecule

For a diatomic molecule, there is only one vibrational mode, so there will be only a single set of vibrational wavefunctions with associated energies for this system. For polytatomic molecules, there will be a set of wavefunctions with associated energy associated with each vibrational mode.

The Hamiltonian operator, the general quantum mechanical operator for energy, includes both a kinetic energy term, $\hat{T}$, and a potential energy term, $\hat{V}$.

$$\hat{H} = \hat{T} + \hat{V} \tag{5.6.1}$$

For the free particle and the particle in a box, the potential energy term used in the Hamiltonian was zero. As shown in Equation 5.6.1, the classical expression for the energy of a harmonic oscillator includes both a kinetic energy term and the harmonic potential energy term. Transforming this equation into the corresponding Hamiltonian operator gives,

$$\hat{H}(q) = \frac{1}{2\mu}\hat{P}_q^2 + \frac{k}{2}\hat{q}^2 \tag{5.6.2}$$

where $\hat{q}$ is the operator for the length of the normal coordinate, and $\hat{P}_q$ is the momentum operator associated with the normal coordinate. μ is an effective (reduced) mass, and k is an effective force constant, and these quantities will be different for each of the normal modes (vibrations).

Substituting the definitions for the operators yields

$$\hat{H}(q) = -\frac{\hbar^2}{2\mu}\frac{d^2}{dq^2} + \frac{k}{2}q^2 \tag{5.6.3}$$

since the operator for position or displacement is just the position or displacement. The time-independent Schrödinger Equation then becomes

$$-\frac{\hbar^2}{2\mu}\frac{d^2\psi_v(q)}{dq^2} + \frac{k}{2}q^2\psi_v(q) = E_v\psi_v(q) \tag{5.6.4}$$

or upon rearranging

$$\frac{d^2\psi_v(q)}{dq^2} + \frac{2\mu}{\hbar^2}\left(E_v - \frac{k}{2}q^2\right)\psi_v(q) = 0 \tag{5.6.5}$$

This differential equation is not straightforward to solve. Rather than fully develop the details of the solution, we will outline the method used because it represents a common strategy for solving differential equations. The steps taken to solve Equation 5.6.5 are to simplify the equation by collecting constants in the parameter β

$$\beta^2 = \frac{\hbar}{\sqrt{\mu k}} \tag{5.6.6}$$

and then changing the variable from q to x where

$$x = \frac{q}{\beta} \tag{5.6.7}$$

so that

$$\frac{d^2}{dq^2} = \frac{1}{\beta^2}\frac{d^2}{dx^2} \tag{5.6.8}$$

After substituting Equations 5.6.6 and 5.6.8 into Equation 5.6.5, the differential equation for the harmonic oscillator becomes

$$\frac{d^2\psi_v(x)}{dx^2} + \left(\frac{2\mu\beta^2 E_v}{\hbar^2} - x^2\right)\psi_v(x) = 0 \tag{5.6.9}$$

? Exercise 5.6.1

Make the substitutions given in Equations 5.6.6 and 5.6.8 into Equation 5.6.5 to get Equation 5.6.9.

Solving for the Quantum Wavefunctions

A common strategy for solving differential equations, which is employed here, is to find a solution that is valid for large values of the variable and then develop the complete solution as a product of this asymptotic solution and a power series. Since the potential energy approaches infinity as x and the coordinate q approach infinity, the wavefunctions must approach zero (this is, the wavefunctions must **converge to zero**):

$$\lim_{x\to\infty} \psi_v(x) = 0 \tag{5.6.10}$$

The function that has this property and satisfies the differential equation for large values of x is the exponential function, i.e.,

$$\lim_{x\to\infty} \phi_v(x)\exp\left(\frac{-x^2}{2}\right) = 0 \tag{5.6.11}$$

where the wavefunctions are

$$\psi_v(x) \propto \phi_v(x)\exp\left(\frac{-x^2}{2}\right)$$

The general expression for a power series of $\phi_v(x)$ is

$$\begin{aligned}\phi_v(x) &= \sum_{n=0}^{\infty} c_n(v)x^n \\ &= c_n(v)x + c_n(v)x^2 + c_n(v)x^3 + \cdots\end{aligned}$$

which can be truncated after the first term, after the second term, after the third term, etc. to produce a set of polynomials. There is one polynomial for each value of v where v can be equal to any integer value including zero.

$$\sum_{n=0}^{v} c_n x^n \tag{5.6.12}$$

Each of the truncations of the power series in Equation 5.6.12 can be multiplied by the exponential function in Equation 5.6.11 to create a family of valid solutions to the differential equation in Equation 5.6.9.

$$\psi_v(x) = \sum_{n=0}^{v} c_n x^n \exp\left(\frac{-x^2}{2}\right) \tag{5.6.13}$$

? Exercise 5.6.2

Write the first four polynomials, $v = 0$ to $v = 1$, $v = 12$, $v = 13$, $v = 14$ for Equation 5.6.12 and use suitable software to prepare plots of these polynomials. Identify the curves in the plots.

? Exercise 5.6.3

Confirm that the wavefunction given by Equation 5.6.13 is a solution to the harmonic oscillator Schrödinger Equation in Equation 5.6.9 for $v = 0$ and $v = 1$.

Hermite Polynomials

While polynomials in general approach ∞ (or $-\infty$) as x approaches ∞, the decreasing exponential term overpowers the polynomial term so that the overall wavefunction exhibits the desired approach to zero at large values of x or $-x$. The exact forms of

polynomials that solve Equation 5.6.9 are the **Hermite polynomials**, which are standard mathematical functions known from the work of Charles Hermite. The first eight Hermite polynomials, $H_v(x)$, are given below.

- $H_0 = 1$
- $H_1 = 2x$
- $H_2 = -2 + 4x^2$
- $H_3 = -12x + 8x^3$
- $H_4 = 12 - 48x^2 + 16x^4$
- $H_5 = 120x - 160x^3 + 32x^5$
- $H_6 = -120 + 720x^2 - 480x^4 + 64x^6$
- $H_7 = -1680x + 3360x^3 - 1344x^5 + 128x^7$

The first six Hermite polynomials are plotted in Figure 5.6.1 . Hermite polynomials will be discussed in more detail in the following Section.

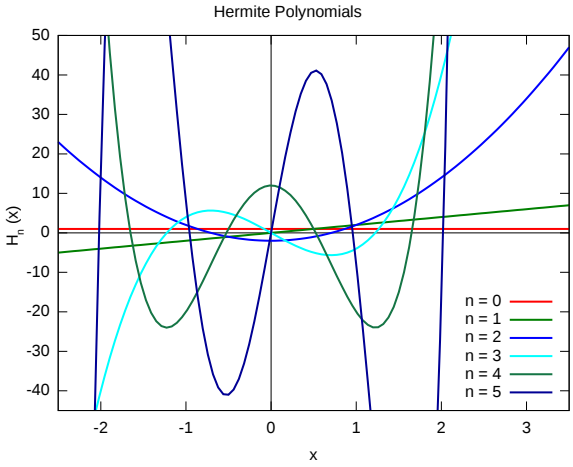


Figure 5.6.1 : The first six Hermite polynomials $H_n(x)$. (CC BY-SA 3.0 Unported; Alessio Damato, Vulpecula and others via Wikipedia)

? Exercise 5.6.4

Determine the units of β and the units of x in the Hermite polynomials.

Because of the association of the wavefunction with a probability density, it is necessary for the wavefunction to include a normalization constant, N_v.

$$N_v = \frac{1}{(2^v v! \sqrt{\pi})^{1/2}} \tag{5.6.14}$$

The final form of the harmonic oscillator wavefunctions is thus

$$\psi_v(x) = N_v H_v(x) e^{-x^2/2} \tag{5.6.15}$$

Alternative and More Common Formulation of Harmonic Oscillator Wavefunctions

The harmonic oscillator wavefunctions are often written in terms of Q, the unscaled displacement coordinate (Equation 5.6.7) and a different constant α:

$$\alpha = 1/\sqrt{\beta} = \sqrt{\frac{k\mu}{\hbar^2}}$$

so Equation 5.6.15 becomes

$$\psi_v(x) = N_v'' H_v(\sqrt{\alpha} Q) e^{-\alpha Q^2/2}$$

with a slightly different normalization constant

$$N_v'' = \sqrt{\frac{1}{2^v v!}} \left(\frac{\alpha}{\pi}\right)^{1/4}$$

Exercise 5.6.5

Compute the normalization factor for $\psi_v(x)$ where $v = 0$ and $v = 4$. What is the purpose of N_v?

The energy eigenvalues for a quantum mechanical oscillator also are obtained by solving the Schrödinger equation. The energies are restricted to discrete values

$$E_v = \left(v + \frac{1}{2}\right) \hbar\omega \tag{5.6.16}$$

with $v = 0, 1, 2, 3, \cdots$.

The energies depend both on the quantum number, v, and the oscillator frequency

$$\omega = \sqrt{\frac{k}{\mu}}$$

which in turn depends on the spring constant k and the reduced mass of the vibration μ.

Exercise 5.6.6

Determine the energy for the first ten harmonic oscillator energy levels in terms of $\hbar\omega$. Sketch an energy level diagram of these energies.

1. What insights do you gain from Equation 5.6.16, your calculations, and your diagram?
2. Is it possible to have a molecule that is not vibrating?
3. In terms of $\hbar\omega$, what is the energy of the photon required to cause a transition from one vibrational state to the next higher one?
4. If a transition from energy level $v = 9$ to $v = 10$ were observed in a spectrum, where would that spectral line appear relative to the one for the transition from level $v = 0$ to $v = 1$?
5. If a vibrational transition is observed at 3000 cm-1 in an infrared spectrum, what is the value of $\hbar\omega$ for the normal mode?
6. Identify all the possible meanings of $\Delta E = h\nu$ and the definition of the frequency, ν, in each case.

The normalized wavefunctions for the first four states of the harmonic oscillator are shown in Figure 5.6.2 , and the corresponding probability densities are shown in Figure 5.6.3 . You should remember the mathematical and graphical forms of the first few harmonic oscillator wavefunctions, and the correlation of v with E_v. The number of nodes in the wavefunction will help you to remember these characteristics. Also note that the functions fall off exponentially and that the symmetry alternates. For v equal to an even number, ψ_v is gerade; for v equal to an odd number, ψ_v is ungerade.

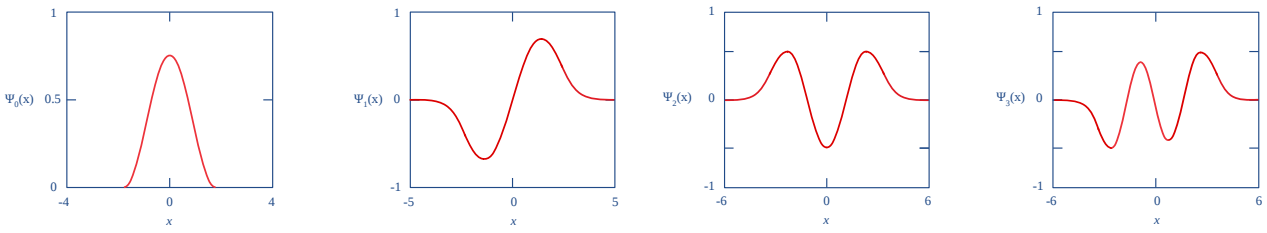


Figure 5.6.2 : The harmonic oscillator wavefunctions describing the four lowest energy states.

Exercise 5.6.7

Write a few sentences describing and comparing the plots in Figure 5.6.2 . How many nodes are there as a function of v? Do the wavefunctions converge at extreme displacement? Where is the most likely displacements for the oscillator to be found?

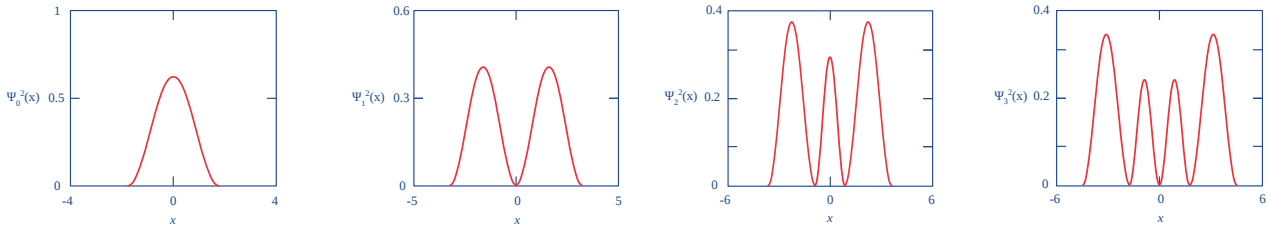


Figure 5.6.3 : The probability densities for the four lowest energy states of the harmonic oscillator.

Exercise 5.6.8

Explain how Figure 5.6.3 is related to Figure 5.6.2 . Explain the physical significance of the plots in Figure 5.6.3 in terms of the magnitude of the normal coordinate Q. Couch your discussion in terms of the HCl molecule. How would you describe the location of the atoms in each of the states? How does the oscillator position correspond to the energy of a particular level?

Exercise 5.6.9

Plot the probability density for energy level 10 of the harmonic oscillator. How many nodes are present? Plot the probability density for energy level 20. Compare the plot for level 20 with that of level 10 and level 1. Compare these quantum mechanical probability distributions to those expected for a classical oscillator. What conclusion can you draw about the probability of the location of the oscillator and the length of a chemical bond in a vibrating molecule? Extend your analysis to include a very high level, like level 50.

In completing Exercise 5.6.9 , you should have noticed that as the quantum number increases and becomes very large, the probability distribution approaches that of a classical oscillator. This observation is very general. It was first noticed by Bohr, and is called the Bohr Correspondence Principle. This principle states that classical behavior is approached in the limit of large values for a quantum number. A classical oscillator is most likely to be found in the region of space where its velocity is the smallest. This situation is similar to walking through one room and running through another. In which room do you spend more time? Where is it more likely that you will be found?

Examination of the quantum mechanical wavefunction for the lowest-energy state reveals that the wavefunction $\psi_0(x)$ extends beyond the classical limit (i.e., outside of the harmonic oscillator well, albeit slightly). Higher energy states have higher total energies, so the classical limits to the amplitude of the displacement will be larger for these states.

Tunneling in the Quantum Harmonic Oscillator

The observation that the wavefunctions are not zero at the classical limit means that the quantum mechanical oscillator has a finite probability of having a displacement that is larger than what is classically possible. The oscillator can be in a region of space where the potential energy is greater than the total energy. Classically, when the potential energy equals the total energy, the kinetic energy and the velocity are zero, and the oscillator cannot pass this point. A quantum mechanical oscillator, however, has a finite probability of passing this point. For a molecular vibration, this property means that the amplitude of the vibration is larger than what it would be in a classical picture. In some situations, a larger amplitude vibration could enhance the chemical reactivity of a molecule.

Exercise 5.6.10

Plot the probability density for $v = 0$ and $v = 1$ states. Mark the classical limits on each of the plots, since the limits are different because the total energy is different for $v = 0$ and $v = 1$. Shade in the regions of the probability densities that extend beyond the classical limit.

We should be able to calculate the probability that the quantum mechanical harmonic oscillator is in the classically forbidden region for the lowest energy state of the harmonic oscillator, the state with $v = 0$. The classically forbidden region is shown by the shading of the regions beyond Q_0 in the graph you constructed for Exercise 5.6.3 . The area of this shaded region gives the probability that the bond oscillation will extend into the forbidden region (Figure 5.6.3). To calculate this probability, we use

$$P[\text{forbidden}] = 1 - P[\text{allowed}] \tag{5.6.17}$$

because the integral from 0 to Q_0 for the allowed region can be found in integral tables and the integral from Q_0 to ∞ cannot. The form of the integral, $P[\text{allowed}]$, to evaluate is

$$P[\text{allowed}] = 2\int_0^{Q_0} \psi_0^*(Q)\psi_0(Q)dQ \tag{5.6.18}$$

The factor 2 appears in Equation 5.6.18 from the chancing the limits of integration from $-Q_0$ to $+Q_0$ into 0 to $+Q_0$; we can do this since the integrand is an even function, i. e., $f(-x) = f(x)$. To evaluate the integral in Equation 5.6.18, use the wavefunction and do the integration in terms of x. Recall that for $v = 0$, $Q = Q_0$ corresponds to $x = 1$. Including the normalization constant, Equation 5.6.18 produces

$$P[\text{allowed}] = \frac{2}{\sqrt{\pi}}\int_0^1 \exp(-x^2)dx \tag{5.6.19}$$

The integral in Equation 5.6.19 is called an error function (ERF) and can only be evaluated numerically. Values can be found in books of mathematical tables. When the limit of integration is 1, ERF(l) = 0.843 and P[forbidden] = 0.157. This result means that the quantum mechanical oscillator can be found in the forbidden region 16% of the time. This effect is substantial and leads to the phenomenon called *quantum mechanical tunneling*.

? Exercise 5.6.11

Numerically verify that $P[\text{allowed}]$ in Equation 5.6.19 equals 0.843. To obtain a value for the integral do not use symbolic integration or symbolic equals.

5.6: The Harmonic-Oscillator Wavefunctions Involve Hermite Polynomials is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.6: The Harmonic Oscillator Wavefunctions involve Hermite Polynomials** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

5.7: Hermite Polynomials are either Even or Odd Functions

Learning Objectives

- Understand key properties of the Hermite polynomials including orthogonality and symmetry.
- Be proficient at using symmetries of integrands to quickly solve integrals.

Hermite polynomials were defined by Laplace (1810) though in scarcely recognizable form, and studied in detail by Chebyshev (1859). Chebyshev's work was overlooked and they were named later after Charles Hermite who wrote on the polynomials in 1864 describing them as new. They were consequently not new although in later 1865 papers Hermite was the first to define the multidimensional polynomials. The first six Hermite polynomial are plotted in Figure 5.7.1 .

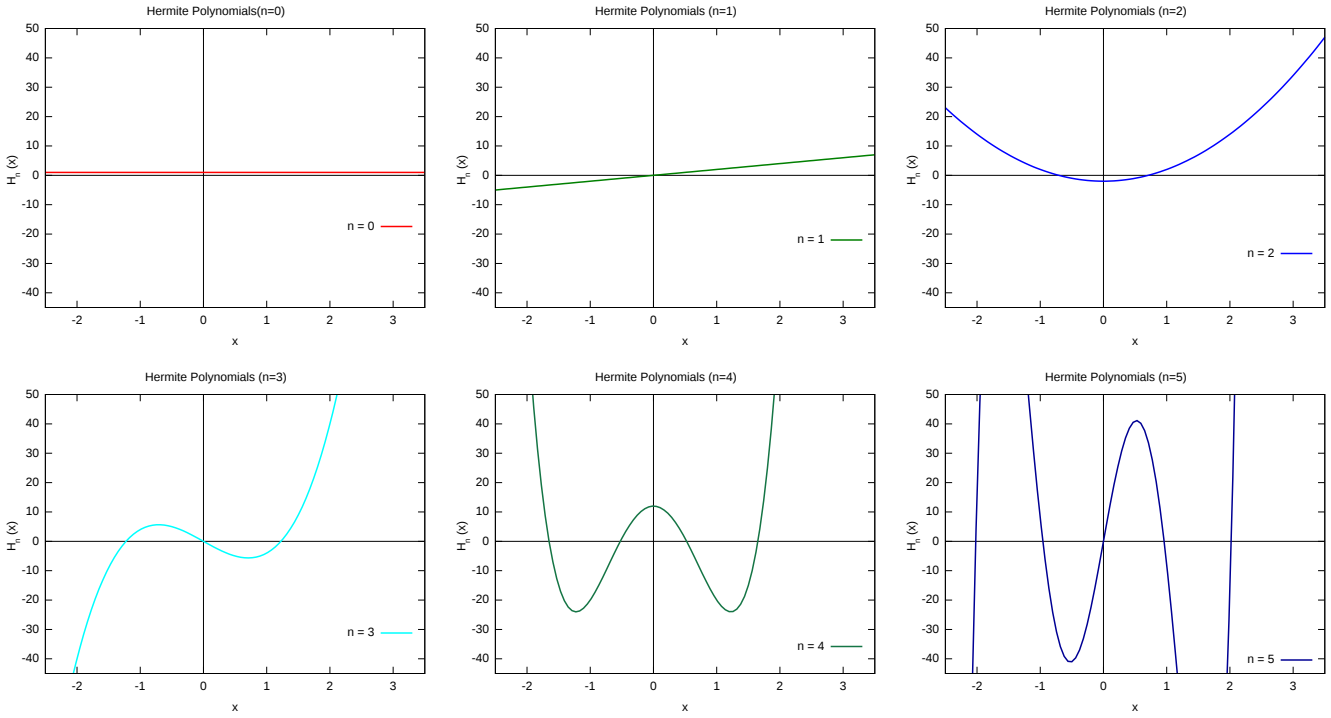


Figure 5.7.1 : The first six Hermite polynomials $H_n(x)$. (CC BY-SA 3.0 Unported; Alessio Damato, Vulpecula and others via Wikipedia)

Generating Formula

Any Hermite polynomial $H_n(x)$ can be generated from a previous one $H_{n-1}(x)$ via the following using the recurrence relation

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x). \tag{5.7.1}$$

Hermite Polynomials are Symmetric

Let $f(x)$ be a real-valued function of a real variable.

- Then f is **even** if the following equation holds for all x and -x in the domain of f

$$f(x) = f(-x)$$

- Then f is **odd** if the following equation holds for all x and -x in the domain of f

$$-f(x) = f(-x)$$

Even and odd are terms used to describe particularly well-behaved functions. An *even* function is symmetric about the y-axis (Figure 5.7.2 ; left). That is, if we reflect the graph of the function in the y-axis, then it does not change. Formally, we say that f is

even if, for all x and $-x$ in the domain of f, we have

$$f(-x) = f(x)$$

Two examples of even functions are $f(x) = x^2$ and $f(x) = \cos x$.

An *odd* function has rotational symmetry of order two about the origin (Figure 5.7.2 ; middle). That is, if we rotate the graph of the function 180° about the origin, then it does not change. Formally, we say that ff is odd if, for all x and $-x$ in the domain of f, we have

$$f(-x) = -f(x)$$

Examples of odd functions are $f(x) = x^3$ and $f(x) = \sin x$.

Naturally, not all functions can be classified as even or odd. For example $f = x^3 + 1$ shown in the right side of Figure 5.7.2 , is neither.

> You can also think of these properties as symmetry conditions at the origin. More symmetries in 3D space are discussed in Group Theory.

Without proof, we can identify several key features involving multiplication properties of even and odd functions:

- The product of two **even** functions is an **even** function.
- The product of two **odd** functions is an **even** function.
- The product of an **even** function and an **odd** function is an odd function.

This can be shown graphically as a product table like that in Table 5.7.1 .

Table 5.7.1 : Product table of 1D Functions

Product table	Odd Function (anti-symmetric)	Even Function (symmetric)	No symmetry (neither)
Odd Function (anti-symmetric)	Even Function (symmetric)	Odd Function (anti-symmetric)	who knows
Even Function (symmetric)	Odd Function (anti-symmetric)	Even Function (symmetric)	who knows
No symmetry (neither)	who knows	who knows	who knows

Notice that the Hermite polynomials in Figure 5.7.1 oscillate from even to odd. We can take advantage of that aspect in our calculation of Harmonic Oscillator wavefunctions. Hermite Polynomial is an even or odd function depends on its degree n. Based on

$$H_n(-x) = (-1)^n H_n(x) \tag{5.7.2}$$

- $H_n(x)$ is an even function, when n is even.
- $H_n(x)$ is an odd function, when n is odd.

Integration over Symmetric Functions

You often consider integrals of the form

$$I = \int_{-a}^{a} f(x)\,\mathrm{d}x$$

If f is odd or even, then sometimes you can make solving this integral easier. For example, we can rewrite that integral in the following way:

$$I = \int_{-a}^{a} f(x)\,\mathrm{d}x = \int_{-a}^{0} f(x)\,\mathrm{d}x + \int_{0}^{a} f(x)\,\mathrm{d}x \tag{5.7.3}$$

$$= \int_{0}^{a} f(-x)\,\mathrm{d}x + \int_{0}^{a} f(x)\,\mathrm{d}x \tag{5.7.4}$$

For an even function, we have $f(-x) = f(x)$ and Equation 5.7.4 can be simplified

$$\begin{aligned} I &= \int_0^a f(-x)\,dx + \int_0^a f(x)\,dx \\ &= 2\int_0^a f(x)\,dx \end{aligned}$$

For an odd function, we have $f(-x) = -f(x)$ and Equation 5.7.4 can be simplified

$$\begin{aligned} I &= -\int_0^a f(x)\,dx + \int_0^a f(x)\,dx \\ &= 0 \end{aligned}$$

That's what it means to simplify the integration: the integral of an odd or even function over the interval $[-L, L]$ can be put into a nicer form (and sometimes we can see that it vanishes without ever computing an integral).

Example 5.7.1

Technically, evaluating the orthogonality of Hermite polynomials requires integrating over the $\exp(-x^2)$ weight function (Equations 5.7.5 and 5.7.6).

Solution

For the Hermite polynomials $H_n(x)$, the relevant inner product (using Dirac Notation)

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\exp(-x^2)\,dx$$

While the $H_2(x)H_3(x)$ product is indeed an odd function (Table 5.7.1), while $exp(-x^2)$ is even. Their product is odd, and thus $\langle f, g\rangle$ certainly ought to be zero.

Symmetry is an important aspect of quantum mechanics and mathematics, especially in calculating integrals. Using this symmetry, integrals can be identified to be equal to zero without explicitly solving them. For example, the integral of an odd integrand over all possible values will **always** be zero irrespective of the exact nature of the function:

$$\int_{-\infty}^{\infty} f(x)\,dx = 0$$

This simplifies calculations greatly as demonstrated in the following chapters.

Hermite Polynomials are Orthogonal

Hermite polynomials $H_n(x)$ are *n*th-degree polynomials for *n* = 0, 1, 2, 3 and form an orthogonal set of functions for the weight function $e^{-x^2/2}$. The exact relation is:

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)e^{-x^2/2}dx = 0 \tag{5.7.5}$$

if $m \neq n$ and

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)e^{-x^2/2}dx = 2^n n!\sqrt{\pi} \tag{5.7.6}$$

if $m = n$.

This will not be proved, but can the demonstrated using any of the Hermite polynomials listed in the previous section. The orthogonality property becomes important when solving the Harmonic oscillator problems. Note that the integral of Equation 5.7.6 is important for normalizing the quantum harmonic oscillator wavefunctions discussed in last Section.

Example 5.7.2 : Hermite Polynomials are Orthogonal

Demonstrate that $H_2(x)$ and $H_3(x)$ are orthogonal.

Solution

We need to confirm

$$\int_{-\infty}^{\infty} H_2(x)H_3(x)dx = 0$$

or when substituted

$$\int_{-\infty}^{\infty} (4x^2-2)(8x^3-12x)dx = 0$$

because it says I need to show it's orthogonal on $[-\infty,\infty]$ or we can just evaluate it on a finite interval $[-L,L]$, where L is a constant.

$$\begin{aligned}\int_{-L}^{L}(4x^2-2)(8x^3-12x)dx &= 8\left(\frac{2x^6}{3}-2x^4+\frac{3x^2}{2}\right)\Bigg|_{-L}^{L}\\ &= 8\left(\frac{2L^6}{3}-2L^4+\frac{3L^2}{2}\right)-8\left(\frac{2(-L)^6}{3}-2(-L)^4+\frac{3(-L)^2}{2}\right)\\ &= 0.\end{aligned}$$

Concluding

Hermite polynomials are a component in the harmonic oscillator wavefunction that dictates the symmetry of the wavefunctions. If your integration interval is symmetric around 0, then the integral over *any* integrable odd function is zero, no exception. Therefore as soon as you've found that your integrand is odd and your integration interval is symmetric, you're done. Also, for general functions, if you can easily split them into even and odd parts, you only have to consider the integral over the even part for symmetric integration intervals.

Another important property is that the product of two even or of two odd functions is even, and the product of an even and an odd function is odd. For example, if ff is even, $x \mapsto f(x)\sin(x)$ is odd, and therefore the integral over it is zero (provided it is well defined).

5.7: Hermite Polynomials are either Even or Odd Functions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.7: Hermite Polynomials are either Even or Odd Functions** by Delmar Larsen is licensed CC BY-SA 4.0.

5.8: The Energy Levels of a Rigid Rotor

Learning Objectives

- Compare the classical and quantum rigid rotor in three dimensions
- Demonstrate how to use the Separation of Variable technique to solve the 3D rigid rotor Schrödinger Equation
- Identify and interpret the two quantum numbers for a 3D quantum rigid rotor including the range of allowed values
- Describe the wavefunctions of the 3D quantum rigid rotor in terms of nodes, average displacements and most probable displacements
- Describe the energies of the 3D quantum rigid rotor in terms of values and degeneracies

Rigid rotor means when the distance between particles do not change as they rotate. A rigid rotor only approximates a rotating diatomic molecular if vibration is ignored.

The Classical Rigid Rotor in 3D

The **rigid rotor** is a mechanical model that is used to explain rotating systems. The linear rigid rotor model consists of two point masses located at fixed distances from their center of mass. The fixed distance between the two masses and the values of the masses are the only characteristics of the rigid model. However, for many actual diatomics this model is too restrictive since distances are usually not completely fixed and corrections on the rigid model can be made to compensate for small variations in the distance. Even in such a case the rigid rotor model is a useful model system to master.

For a rigid rotor, the total energy is the sum of kinetic (T) and potential (V) energies

$$E_{tot} = T + V \tag{5.8.1}$$

The potential energy, V, is set to 0 because the distance between particles does not change within the rigid rotor approximation. However, In reality, $V \neq 0$ because even though the average distance between particles does not change, the particles still vibrate. The rigid rotor approximation greatly simplifys our discussion.

Since $V = 0$ then $E_{tot} = T$ and we can also say that:

$$T = \frac{1}{2}\sum m_i v_i^2 \tag{5.8.2}$$

However, we have to determine v_i in terms of rotation since we are dealing with rotation motion. Since,

$$\omega = \frac{v}{r} \tag{5.8.3}$$

where ω is the **angular velocity**, we can say that:

$$v_i = \omega X r_i \tag{5.8.4}$$

Thus we can rewrite Equation 5.8.2 as:

$$T = \frac{1}{2}\sum m_i v_i \left(\omega X r_i\right) \tag{5.8.5}$$

Since ω is a scalar constant, we can rewrite Equation 5.8.5 as:

$$T = \frac{\omega}{2}\sum m_i \left(v_i X r_i\right) = \frac{\omega}{2}\sum l_i = \omega\frac{L}{2} \tag{5.8.6}$$

where l_i is the **angular momentum** of the ith particle, and L is the angular momentum of the *entire system*. Also, we know from physics that,

$$L = I\omega \tag{5.8.7}$$

where I is the moment of inertia of the rigid body relative to the axis of rotation. We can rewrite Equation 5.8.2 as

$$T = \omega \frac{I\omega}{2} = \frac{1}{2} I\omega^2 \tag{5.8.8}$$

Equation 5.8.8 shows that the energy of the rigid rotor scales with increasing angular frequency (i.e., the faster is rotates) and with increasing moment of inertia (i.e, the inertial resistance to rotation). Also, as expected, the classical rotational energy is not quantized (i.e., all possible rotational frequencies are possible).

The Quantum Rigid Rotor in 3D

It is convenient to discuss rotation with in the spherical coordinate system rather than the Cartesian system (Figure 5.8.1).

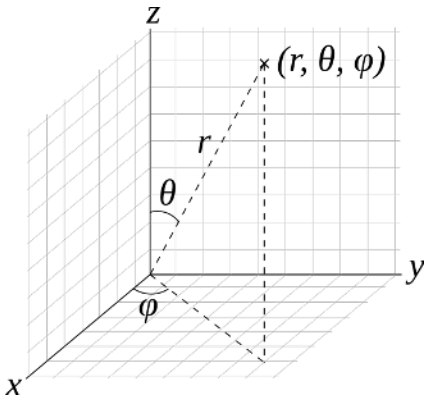


Figure 5.8.1 : The spherical coordinate system in terms of the Cartesian system. (Public Domain; Andeggs via Wikipedia).

To solve the Schrödinger equation for the rigid rotor, we will separate the variables and form single-variable equations that can be solved independently. Only two variables θ and φ are required in the rigid rotor model because the bond length, r, is taken to be the constant r_0. We first write the rigid rotor wavefunctions as the product of a theta-function depending only on θ and a phi-function depending only on φ

$$|\psi(\theta,\varphi)\rangle = |\Theta(\theta)\Phi(\varphi)\rangle \tag{5.8.9}$$

We then substitute the product wavefunction and the Hamiltonian written in spherical coordinates into the Schrödinger Equation 5.8.10

$$\hat{H}|\Theta(\theta)\Phi(\varphi)\rangle = E|\Theta(\theta)\Phi(\varphi)\rangle \tag{5.8.10}$$

to obtain

$$-\frac{\hbar^2}{2\mu r_0^2}\left[\frac{\partial}{\partial r_0} r_0^2 \frac{\partial}{\partial r_0} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]|\Theta(\theta)\Phi(\varphi)\rangle = E|\Theta(\theta)\Phi(\varphi)\rangle \tag{5.8.11}$$

Since $r = r_0$ is constant for the **rigid rotor** and does not appear as a variable in the functions, the partial derivatives with respect to r are zero; i.e. the functions do not change with respect to r. We also can substitute the symbol I for the moment of inertia, μr_0^2 in the denominator of the left hand side of Equation 5.8.11, to give

$$-\frac{\hbar^2}{2I}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]|\Theta(\theta)\Phi(\varphi)\rangle = E|\Theta(\theta)\Phi(\varphi)\rangle \tag{5.8.12}$$

To begin the process of the Separating of Variables technique, multiply each side of Equation 5.8.12 by $\frac{2I}{\hbar^2}$ and $\frac{-\sin^2\theta}{\Theta(\theta)\Phi(\varphi)}$ to give

$$\frac{1}{\Theta(\theta)\psi(\varphi)}\left[\sin\theta\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{\partial^2}{\partial\varphi^2}\right]\Theta(\theta)\Phi(\varphi) = \frac{-2IE\sin^2\theta}{\hbar^2} \tag{5.8.13}$$

Simplify the appearance of the right-hand side of Equation 5.8.13 by defining a parameter λ:

$$\lambda = \frac{2IE}{\hbar^2}. \tag{5.8.14}$$

Note that this λ has no connection to a wavelength; it is merely being used as an algebraic symbol for the combination of constants shown in Equation 5.8.14.

Inserting λ, evaluating partial derivatives, and rearranging Equation 5.8.13 produces

$$\frac{1}{\Theta(\theta)}\left[\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\Theta(\theta)+\left(\lambda\sin^2\theta\right)\Theta(\theta)\right]=-\frac{1}{\Phi(\varphi)}\frac{\partial^2}{\partial\varphi^2}\Phi(\varphi) \tag{5.8.15}$$

? Exercise 5.8.1

Carry out the steps leading from Equation 5.8.13 to Equation 5.8.15. Keep in mind that, if y is not a function of x,

$$\frac{dy}{dx}=y\frac{d}{dx}$$

Equation 5.8.15 says that the function on the left, depending only on the variable θ, always equals the function on the right, depending only on the variable φ, for all values of θ and φ. The only way two different functions of independent variables can be equal for all values of the variables is if both functions are equal to a constant (review separation of variables). We call this constant m_J^2 because soon we will need the square root of it. The two differential equations to solve are the θ-equation

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{d}{d\theta}\right)\Theta(\theta)+\left(\lambda\sin^2\theta-m_J^2\right)\Theta(\theta)=0 \tag{5.8.16}$$

and the φ-equation

$$\frac{d^2}{d\varphi^2}\Phi(\varphi)+m_J^2\Phi(\varphi)=0 \tag{5.8.17}$$

The partial derivatives have been replaced by total derivatives because only a single variable is involved in each equation.

> *Often m_J is referred to as just m for convenience.*

? Exercise 5.8.2

Show how Equations 5.8.16 and 5.8.17 are obtained from Equation 5.8.15.

Solving the φ Equation

The φ-equation is similar to the Schrödinger Equation for the free particle. Since we already solved this previously, we immediately write the solutions:

$$\Phi_m(\varphi)=Ne^{\pm im_J\varphi} \tag{5.8.18}$$

where we introduce the number m to track how many wavelengths of the wavefunction occur around one rotation (similar to the wavelength description of the Bohr atom).

? Exercise 5.8.3

Substitute Equation 5.8.18 into Equation 5.8.17 to show that it is a solution to that differential equation.

The normalization condition, Equation 5.8.19 is used to find a value for N that satisfies Equation 5.8.18.

$$\int_0^{2\pi}\Phi^*(\varphi)\Phi(\varphi)d\varphi=1 \tag{5.8.19}$$

The range of the integral is only from 0 to 2π because the angle φ specifies the position of the internuclear axis relative to the x-axis of the coordinate system and angles greater than 2π do not specify additional new positions.

? Exercise 5.8.4

Use the normalization condition in Equation 5.8.19 to demonstrate that $N=1/\sqrt{2\pi}$.

Values for m are found by using a **cyclic boundary condition**. The cyclic boundary condition means that since φ and $\varphi + 2\varphi$ refer to the same point in three-dimensional space, $\Phi(\varphi)$ must equal $\Phi(\varphi + 2\pi)$, i.e.

$$e^{im_J\varphi} = e^{im_J(\varphi+2\pi)} \tag{5.8.20}$$

$$= e^{im_J\varphi}e^{im_J2\pi} \tag{5.8.21}$$

For the equality in Equation 5.8.21 to hold, $e^{im_J2\pi}$ must equal 1, which is true only when

$$m_J = \cdots, -3, -2, -1, 0, 1, 2, 3, \cdots \tag{5.8.22}$$

In other words m_J can equal **any positive or negative integer or zero.**

? Exercise 5.8.5 : Cyclic Boundary Conditions

Use Euler's Formula to show that $e^{im_J2\pi}$ equals 1 for m_J equal to zero or any positive or negative integer.

Thus, the Φ function is

$$\Phi_{m_J}(\varphi) = \sqrt{\frac{1}{2\pi}}e^{\pm im_J\varphi}$$

with

$$m_J = 0, \pm1, \pm2, \cdots$$

Solving the $\Theta(\theta)$ Equation

Finding the $\Theta(\theta)$ functions that are solutions to the θ-equation (Equation 5.8.16) is a more complicated process. Solutions are found to be a set of power series called *Associated Legendre Functions* (Table M2), which are power series of trigonometric functions, i.e., products and powers of sine and cosine functions. The $\Theta(\theta)$ functions, along with their normalization constants, are shown in the third column of Table 5.8.1 .

Table 5.8.1 : Spherical Harmonic Wavefunctions

m_J	J	$\Theta_J^{m_J}(\theta)$	$\Phi(\varphi)$	$Y_J^{m_J}(\theta,\varphi)$
0	0	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2\pi}}$	$\frac{1}{\sqrt{4\pi}}$
0	1	$\sqrt{\frac{3}{2}}\cos\theta$	$\frac{1}{\sqrt{2\pi}}$	$\sqrt{\frac{3}{4\pi}}\cos\theta$
1	1	$\sqrt{\frac{3}{4}}\sin\theta$	$\frac{1}{\sqrt{2\pi}}e^{i\varphi}$	$\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\varphi}$
-1	1	$\sqrt{\frac{3}{4}}\sin\theta$	$\frac{1}{\sqrt{2\pi}}e^{-i\varphi}$	$\sqrt{\frac{3}{8\pi}}\sin\theta e^{-i\varphi}$
0	2	$\sqrt{\frac{5}{8}}(3\cos^2\theta-1)$	$\frac{1}{\sqrt{2\pi}}$	$\sqrt{\frac{5}{16\pi}}(3\cos^2\theta-1)$
1	2	$\sqrt{\frac{15}{4}}\sin\theta\cos\theta$	$\frac{1}{\sqrt{2\pi}}e^{i\varphi}$	$\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{i\varphi}$
-1	2	$\sqrt{\frac{15}{4}}\sin\theta\cos\theta$	$\frac{1}{\sqrt{2\pi}}e^{-i\varphi}$	$\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{-i\varphi}$
2	2	$\sqrt{\frac{15}{16}}\sin^2\theta$	$\frac{1}{\sqrt{2\pi}}e^{2i\varphi}$	$\sqrt{\frac{15}{32\pi}}\sin^2\theta e^{2i\varphi}$
-2	2	$\sqrt{\frac{15}{16}}\sin^2\theta$	$\frac{1}{\sqrt{2\pi}}e^{2i\varphi}$	$\sqrt{\frac{15}{32\pi}}\sin^2\theta e^{-2i\varphi}$

The solution to the θ-equation requires that λ in Equation 5.8.15 be given by

$$\lambda = J(J+1) \tag{5.8.23}$$

where

$$J \geq |m_J| \tag{5.8.24}$$

J can be 0 or any positive integer greater than or equal to m_J. Each pair of values for the quantum numbers, J and m_J, identifies a rotational state with a wavefunction (Equation 5.8.9) and energy (below). Equation 5.8.24 means that J controls the allowed values of m_J.

> *Each pair of values for the quantum numbers, J and m_J, identifies a rotational state and hence a specific wavefunction with associated energy.*

The combination of Equations 5.8.14 and 5.8.23 reveals that the energy of this system is quantized.

$$E = \frac{\hbar^2 \lambda}{2I} = J(J+1)\frac{\hbar^2}{2I} \tag{5.8.25}$$

Using Equation 5.8.25, you can construct a rotational energy level diagram (Figure 5.8.2). For simplicity, use energy units of $\frac{\hbar^2}{2I}$.

- $J = 0$: The lowest energy state has $J = 0$ and $m_J = 0$. This state has an energy $E_0 = 0$. There is only **one state** with this energy, i.e. one set of quantum numbers, one wavefunction, and one set of properties for the molecule.
- $J = 1$: The next energy level is $J = 1$ with energy $\frac{2\hbar^2}{2I}$. There are **three states** with this energy because m_J can equal +1, 0, or -1. These different states correspond to different orientations of the rotating molecule in space. States with the same energy are said to be degenerate. The degeneracy of an energy level is the number of states with that energy. The degeneracy of the $J = 1$ energy level is 3 because there are three states with the energy $\frac{2\hbar^2}{2I}$.
- $J = 2$: The next energy level is for $J = 2$. The energy is $\frac{6\hbar^2}{2I}$, and there are **five states** with this energy corresponding to m_J $= +2, +1, 0, -1, -2$. The energy level degeneracy is five. Note that the spacing between energy levels increases as J increases. Also note that the degeneracy increases. The degeneracy is always $2J + 1$ because m_J ranges from $+J$ to $-J$ in integer steps, including 0.

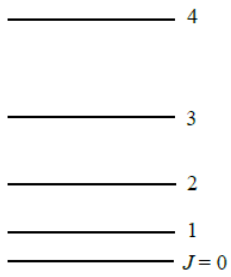


Figure 5.8.2 : Energy spacing for a rigid rotor (in 3D). Notice the energy depends only on J and not depend on m_J. That is, the energy of the rotor is not affected by its orientation.

> *Each allowed energy of rigid rotor is $(2J + 1)$-fold degenerate. Hence, there exist $(2J + 1)$ different wavefunctions with that energy.*

? Exercise 5.8.6

Compute the energy levels for a rotating molecule for $J = 0$ to $J = 5$ using units of $\frac{\hbar^2}{2I}$.

? Exercise 5.8.7

For $J = 0$ to $J = 5$, identify the degeneracy of each energy level and the values of the m_J quantum number that go with each value of the J quantum number. Construct a rotational energy level diagram including $J = 0$ through $J = 5$. Label each level with the appropriate values for the quantum numbers J and m_J. Describe how the spacing between levels varies with increasing J.

Interpretation of Quantum Numbers for a Rigid Rotor

The m_J quantum number reflects the component of the angular momentum along the z direction (and hence is sometimes called the azimuthal quantum number). For a fixed value of J, the different values of m_J reflect the different **directions** the angular momentum vector could be pointing – for large, positive m_J the angular momentum is mostly along +z; if m_J is zero the angular momentum is orthogonal to z. Physically, the energy of the rotation does not depend on the direction, which is reflected in the fact that the energy depends only on J (Equation 5.8.25), which measures the length of the vector, not its direction given by m_J.

Example 5.8.7 : Molecular Oxygen

Calculate $J = 0$ to $J = 1$ rotational transition of the O_2 molecule with a bond length of 121 pm.

Solution

$$E = \frac{\hbar^2}{I} = \frac{\hbar^2}{\mu r^2}$$

$$\mu_{O2} = \frac{m_O m_O}{m_O + m_O} = \frac{(15.9994)(15.9994)}{15.9994 + 15.9994} = 7.9997$$

convert from atomic units to kilogram using the conversion: 1 au = 1.66 x 10^{-27} kg. Plug and chug.

$$E = 5.71 \times 10^{-27}\ Joules$$

Spherical Harmonics

A wavefunction that is a solution to the rigid rotor Schrödinger Equation (Equation 5.8.9) can be written as a single function $Y(\theta, \varphi)$, which is called a spherical harmonic function.

$$Y_J^{m_J}(\theta, \varphi) = \Theta_J^{|m_J|}(\theta)\Phi_{m_J}(\varphi) \tag{5.8.26}$$

The spherical harmonic wavefunction is labeled with m_J and J because its functional form depends on both of these quantum numbers. These functions are tabulated above for $J = 0$ through $J = 2$ and for $J = 3$ in the Spherical Harmonics Table (M4) Polar plots of some of the θ-functions are shown in Figure 5.8.3 .

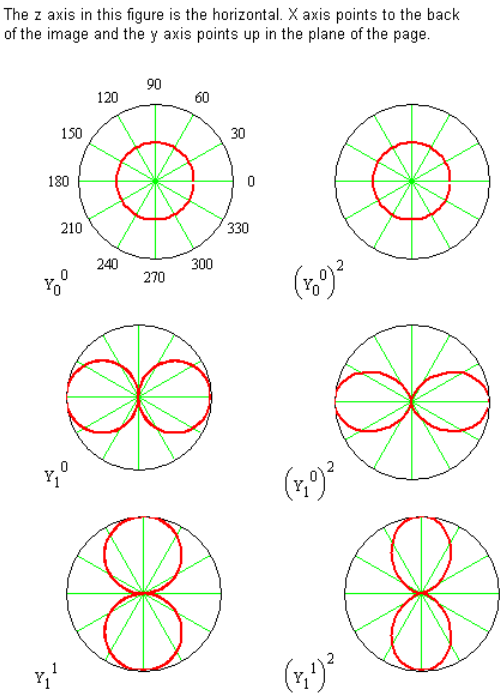


Figure 5.8.3 : Polar plots in which the distance from the center gives the value of the function Y for the indicated angle θ.

The two-dimensional space for a rigid rotor is defined as the surface of a sphere of radius r_0, as shown in Figure 5.8.2 .

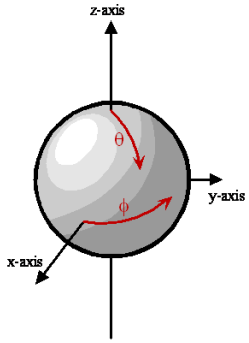


Figure 5.8.4 : Space for a rigid rotor is restricted to the surface of a sphere of radius r_0. The only degrees of freedom are motions along θ or φ on the surface of the sphere.

The probability of finding the internuclear axis at specific coordinates θ_0 and φ_0 within an infinitesimal area ds on this curved surface is given by

$$Pr\left[\theta_0, \varphi_0\right] = Y_J^{m_J*}(\theta_0, \varphi_0) Y_J^{m_J}(\theta_0, \varphi_0) ds \tag{5.8.27}$$

where the area element ds is centered at θ_0 and φ_0.

Within the Copenhagen interpretation of wavefunctions, the absolute square (or modulus squared) of the rigid rotor wavefunction $Y_J^{m_J*}(\theta, \varphi) Y_J^{m_J}(\theta, \varphi)$ gives the probability density for finding the internuclear axis oriented at θ to the z-axis and φ to the x-axis. In spherical coordinates the area element used for integrating θ and φ is

$$ds = \sin\theta \, d\theta \, d\varphi \tag{5.8.28}$$

? Exercise 5.8.8

Use calculus to evaluate the probability of finding the internuclear axis of a molecule described by the $J = 1$, $m_J = 0$ wavefunction somewhere in the region defined by a range in θ of 0° to 45°, and a range in of 0° to 90°. Note that a double

integral will be needed. Sketch this region as a shaded area on Figure 5.8.3 .

Consider the significance of the probability density function by examining the $J = 1$, $m_J = 0$ wavefunction. The Spherical Harmonic for this case is

$$Y_1^0 = \sqrt{\frac{3}{4\pi}} \cos\theta \tag{5.8.29}$$

The polar plot of $(Y_1^0)^2$ is shown in Figure 5.8.3 . For $J = 1$ and $m_J = 0$, the probability of finding the internuclear axis is independent of the angle φ from the x-axis, and greatest for finding the internuclear axis along the z-axis, but there also is a probability for finding it at other values of θ as well. So, although the internuclear axis is not always aligned with the z-axis, the probability is highest for this alignment. Also, since the probability is independent of the angle φ, the internuclear axis can be found in any plane containing the z-axis with equal probability.

The $J = 1$, $m_J = 0$ function is 0 when θ = 90°. Therefore, the entire xy-plane is a **node**. This fact means the probability of finding the internuclear axis in this particular horizontal plane is 0 in contradiction to our classical picture of a rotating molecule. In the classical picture, a molecule rotating in a plane perpendicular to the xy-plane must have the internuclear axis lie in the xy-plane twice every revolution, but the quantum mechanical description says that the probability of being in the xy-plane is zero. This conclusion means that molecules are not rotating in the classical sense, but they still have some, but not all, of the properties associated with classical rotation. The properties they retain are associated with angular momentum.

? Exercise 5.8.9

For each state with $J = 0$ and $J = 1$, use the function form of the Y spherical harmonics and Figure 5.8.1 to determine the most probable orientation of the internuclear axis in a diatomic molecule, i.e., the most probable values for θ and θ.

? Exercise 5.8.10

Write a paragraph describing the information about a rotating molecule that is provided in the polar plot of $Pr[\theta, \theta]$ for the $J = 1$, $m_J = \pm 1$ state in Figure 5.8.1 . Compare this information to the classical picture of a rotating object.

Summary

There are two quantum numbers that describe the quantum behavior of a rigid rotor in three-deminesions: J is the total angular momentum quantum number and m_J is the z-component of the angular momentum. The spherical harmonics called $Y_J^{m_J}$ are functions whose probability $|Y_J^{m_J}|^2$ has the well known shapes of the s, p and d orbitals etc learned in general chemistry.

References

1. Anderson, J.M. Introduction to Quantum Chemistry, 1969, W.A. Benjamin, Inc, pg.91-100.
2. Atkins, Peter and de Paula, Julio. Physical Chemistry for the Life Sciences. New York: W.H. Freeman and Company. p. 515.

Contributors and Attributions

- Ian Huh
- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")
- Mark Tuckerman (New York University)
- MIT OpenCourseWare (Robert Guy Griffin and Troy Van Voorhis)

5.8: The Energy Levels of a Rigid Rotor is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.8: The Energy Levels of a Rigid Rotor** is licensed CC BY-NC-SA 4.0.

5.9: The Rigid Rotator is a Model for a Rotating Diatomic Molecule

Learning Objectives

- Demonstrate how to use the 3D regid rotor to describe a rotating diatomic molecules
- Demonstate how microwave spectroscopy can get used to characterize rotating diatomic molecules
- Interprete a simple microwave spectrum for a diatomic molecule

To develop a description of the rotational states, we will consider the molecule to be a rigid object, i.e. the bond lengths are fixed and the molecule cannot vibrate. This model for rotation is called the rigid-rotor model. It is a good approximation (even though a molecule vibrates as it rotates, and the bonds are elastic rather than rigid) because the amplitude of the vibration is small compared to the bond length.

The rotation of a rigid object in space is very simple to visualize. Pick up any object and rotate it. There are orthogonal rotations about each of the three Cartesian coordinate axes just as there are orthogonal translations in each of the directions in three-dimensional space (Figures 5.9.1 and 5.9.2). These rotations are said to be *orthogonal* because one can not describe a rotation about one axis in terms of rotations about the other axes just as one can not describe a translation along the x-axis in terms of translations along the y- and z-axes. For a linear molecule, the motion around the interatomic axis (x-axis) is not considered a rotation.

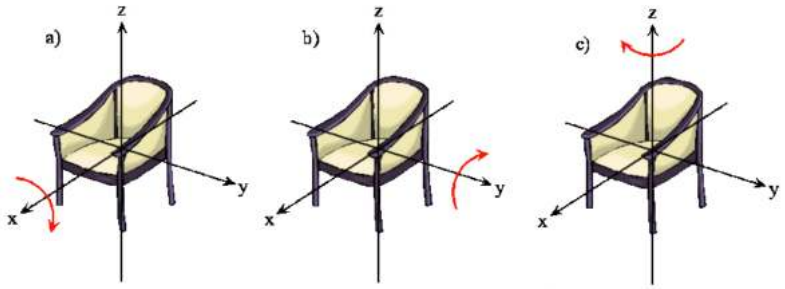


Figure 5.9.1 : Cartesian rotations of a chair. Rotations about the (a) x axis, (b) y axis, and (c) z axis.

In this section we examine the rotational states for a diatomic molecule by comparing the classical interpretation of the angular momentum vector with the probabilistic interpretation of the angular momentum wavefunctions. We want to answer the following types of questions. How do we describe the orientation of a rotating diatomic molecule in space? Is the molecule actually rotating? What properties of the molecule can be physically observed? In what ways does the quantum mechanical description of a rotating molecule differ from the classical image of a rotating molecule?

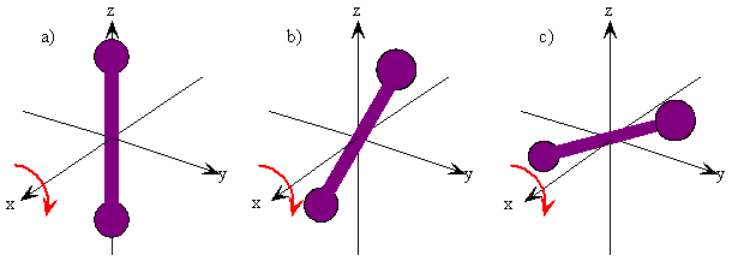


Figure 5.9.2 : Rotation of a diatomic molecule about the x axis.

Introduction to Microwave Spectroscopy

The permanent electric dipole moments of polar molecules can couple to the electric field of electromagnetic radiation. This coupling induces transitions between the rotational states of the molecules. The energies that are associated with these transitions are detected in the far infrared and microwave regions of the spectrum. For example, the microwave spectrum for carbon monoxide spans a frequency range of 100 to 1200 GHz, which corresponds to 3 - 40 cm^{-1}.

The selection rules for the rotational transitions are derived from the transition moment integral by using the spherical harmonic functions and the appropriate dipole moment operator, $\hat{\mu}$.

$$\mu_T = \int Y_{J_f}^{m_f *} \hat{\mu} Y_{J_i}^{m_i} \sin\theta \, d\theta \, d\varphi \tag{5.9.1}$$

or in braket notation

$$\mu_T = \langle Y_{J_f}^{m_f} | \hat{\mu} | Y_{J_i}^{m_i} \rangle \tag{5.9.2}$$

Evaluating the transition moment integral involves a bit of mathematical effort. This evaluation reveals that the transition moment depends on the square of the dipole moment of the molecule, μ^2 and the rotational quantum number, J, of the initial state in the transition,

$$\mu_T = \mu^2 \frac{J+1}{2J+1} \tag{5.9.3}$$

and that the selection rules for rotational transitions are

$$\Delta J = \pm 1 \tag{5.9.4}$$

and

$$\Delta m_J = 0, \pm 1 \tag{5.9.5}$$

A photon is absorbed for $\Delta J = +1$ and emitted for $\Delta J = -1$.

? Exercise 5.9.1

Explain why your microwave oven heats water, but not air. Hint: draw and compare Lewis structures for components of air and for water.

The energies of the Jth rotational levels are given by

$$E_J = J(J+1)\frac{\hbar^2}{2I} \tag{5.9.6}$$

with each Jth energy level having a degeneracy of $2J+1$ due to the different possible m_J values.

Microwave Transition Energies

The transition energies for absorption of radiation are given by

$$E_{photon} = \Delta E \tag{5.9.7}$$
$$= E_f - E_i \tag{5.9.8}$$
$$= h\nu \tag{5.9.9}$$
$$= hc\bar{\nu} \tag{5.9.10}$$

Substituting the relationship for energy (Equation 5.9.6) into Equation 5.9.8 results in

$$E_{photon} = E_f - E_i \tag{5.9.11}$$
$$= J_f(J_f+1)\frac{\hbar^2}{2I} - J_i(J_i+1)\frac{\hbar^2}{2I} \tag{5.9.12}$$

with J_i and J_f representing the rotational quantum numbers of the initial (lower) and final (upper) levels involved in the absorption transition.

Since microwave spectroscopists use frequency units and infrared spectroscopists use wavenumber units when describing rotational spectra and energy levels, both ν and $\bar{\nu}$ are important to calculate. When we add in the constraints imposed by the selection rules to identify possible transitions, J_f in Equation 5.9.12 can be replaced by $J_i + 1$, since the selection rule requires $J_f - J_i = 1$ for the

absorption of a photon (Equation 5.9.4). The equation for absorption transitions (Equation 5.9.12) then can be written in terms of the only the quantum number J_i of the initial state.

$$E_{photon} = h\nu \tag{5.9.13}$$

$$= hc\bar{\nu} \tag{5.9.14}$$

$$= 2(J_i + 1)\frac{\hbar^2}{2I} \tag{5.9.15}$$

Equation 5.9.15 can be rewritten as

$$E_{photon} = 2B(J_i + 1)$$

where B is the **rotational constant** for the molecule and is defined in terms of the *energy* of the absorbed photon

$$B = \frac{\hbar^2}{2I} \tag{5.9.16}$$

Often spectroscopists want to express the rotational constant in terms of *frequency* of the absorbed photon and do so by dividing Equation 5.9.16 by h

$$B(\text{in freq}) = \frac{B}{h} \tag{5.9.17}$$

$$= \frac{h}{8\pi^2\mu r_0^2} \tag{5.9.18}$$

More often, spectroscopists want to express the rotational constant in terms of *wavenumbers* ($\bar{\nu}$) of the absorbed photon by dividing Equation 5.9.16 by hc,

$$\tilde{B} = \frac{B}{hc} = \frac{h}{8\pi^2 c\mu r_0^2} \tag{5.9.19}$$

The rotational constant depends on the distance (R) and the masses of the atoms (via the reduced mass) of the nuclei in the diatomic molecule.

? Exercise 5.9.2

Construct a rotational energy level diagram for $J = 0$, 1, and 2 and add arrows to show all the allowed transitions between states that cause electromagnetic radiation to be absorbed or emitted.

? Exercise 5.9.3

Complete the steps going from Equation 5.9.12 to Equation 5.9.16 and identify the units of B at the end.

? Exercise 5.9.4

Infrared spectroscopists use units of wavenumbers. Rewrite the steps going from Equation 5.9.12 to Equation 5.9.16 to obtain expressions for $h\nu$ and B in units of wavenumbers. Note that to convert B in Hz to B in cm^{-1}, you simply divide the former by c.

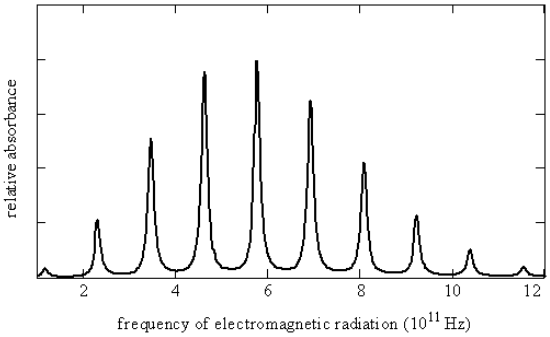


Figure 5.9.3: The rotation spectrum of $^{12}{}^{16}CO$ at 40 K. The peaks are tabulated in Table 5.9.1.

Figure 5.9.3 shows the rotational spectrum as a series of nearly equally spaced lines. The line positions ν_J, line spacings, and the maximum absorption coefficients (γ_{max}, the absorption coefficients associated with the specified line position) for each line in this spectrum are given here in Table 5.9.1.

Table 5.9.1 : Rotational Transitions in $^{12}C^{16}O$ at 40 K

J	ν_J (MHz)	Spacing from previous line (MHz)	γ_{max}
$0 \rightarrow 1$	115,271.21	0	0.0082
$1 \rightarrow 2$	230,538.01	115,266.80	0.0533
$2 \rightarrow 3$	345,795.99	115,257.99	0.1278
$3 \rightarrow 4$	461,040.76	115,244.77	0.1878
$4 \rightarrow 5$	576,267.91	115,227.15	0.1983
$6 \rightarrow 6$	691,473.03	115,205.12	0.1618
$6 \rightarrow 7$	806,651.78	115,178.68	0.1064
$7 \rightarrow 8$	921,799.55	115,147.84	0.0576
$8 \rightarrow 9$	1,036,912.14	115,112.59	0.0262
$9 \rightarrow 10$	1,151,985.08	115,072.94	0.0103

Let's try to reproduce Figure 5.9.3 from the data in Table 5.9.1 by using the quantum theory that we have developed so far. Equation 5.9.19 predicts a pattern of exactly equally spaced lines. The lowest energy transition is between $J_i = 0$ and $J_f = 1$ so the first line in the spectrum appears at a frequency of $2B$. The next transition is from $J_i = 1$ to $J_f = 2$ so the second line appears at $4B$. The spacing of these two lines is $2B$. In fact the spacing of all the lines is $2B$, which is consistent with the experimental data in Table 5.9.1 showing that the lines are very nearly equally spaced. The difference between the first spacing and the last spacing is less than 0.2%.

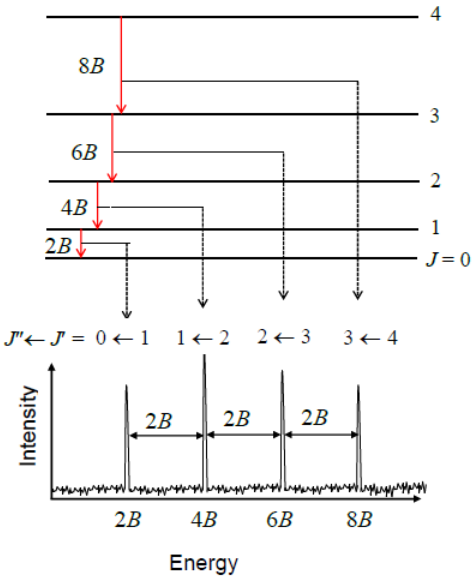


Figure 5.9.4 : Energy levels and line positions calculated in the rigid rotor approximation. (CC BY-SA 3.0; Nrrw via Wikipedia)

Exercise 5.9.5

Use Equation 5.9.19 to prove that the spacing of any two lines in a rotational spectrum is $2B$, i.e. derive:

$$\nu_{J_i+1} - \nu_{J_i} = 2B$$

Example 5.9.1 : Rotation of Sodium Hydride

The molecule $\mathbf{NaH}$ undergoes a rotational transition from $J = 0$ to $J = 1$ when it absorbs a photon of frequency $2.94 \times 10^{11}\ Hz$. What is the equilibrium bond length of the molecule?

Solution

We use $J = 0$ in the formula for the transition frequency

$$\nu = 2B = \frac{\hbar}{2\pi I} = \frac{\hbar}{2\pi\mu R_e^2}$$

Solving for R_e gives

$$R_e = \sqrt{\frac{\hbar}{2\pi\mu\nu}}$$

The reduced mass is given by

$$\begin{aligned}\mu &= \frac{m_{Na}m_H}{m_{Na}+m_H} \\ &= \frac{(22.989)(1.0078)}{22.989+1.0078} \\ &= 0.9655\end{aligned}$$

which is in atomic mass units or relative units. To convert to kilograms, we need the conversion factor $1\ au = 1.66 \times 10^{-27}\ kg$. Multiplying this by 0.9655 gives a reduced mass of $1.603 \times 10^{-27}\ kg$. Substituting in for R_e gives

$$\begin{aligned}R_e &= \sqrt{\frac{(1.055 \times 10^{-34}\ J \cdot s)}{2\pi(1.603 \times 10^{-27}\ kg)(2.94 \times 10^{11}\ Hz)}} \\ &= 1.899 \times 10^{-10}\ m \\ &= 1.89\ \mathring{A}\end{aligned}$$

Example 5.9.6

Use the frequency of the $J = 0$ to $J = 1$ transition observed for carbon monoxide to determine a bond length for $^{12}C^{16}O$.

Solution

- J=0: $v_0 = 115271.21\ \text{MHz}$
- J=1: $v_1 = 230538.01\ \text{MHz}$

$$\begin{aligned}\Delta v &= 230538.01\ \text{MHz} - 115271.21\ \text{MHz} \\ &= 115266.8\ \text{MHz} \\ &= 1.153 \times 10^{11}\ \text{Hz} \\ &= \frac{\hbar}{2\pi\mu R_e^2}\end{aligned}$$

The reduced mass is

$$\begin{aligned}\mu &= \frac{m_C m_O}{m_C + m_O} \\ &= \frac{12.01 \times 16.00}{12.01 + 16.00} \\ &= 6.86\ \text{amu}\end{aligned}$$

Convert to kg

$$6.86 amu \left(\frac{1.661 \times 10^{-27}\ \text{kg}}{12\ \text{amu}}\right) = 1.139 \times 10^{-26}\ \text{kg}$$

$$\begin{aligned}R_e &= \sqrt{\frac{\hbar}{2\pi\mu\Delta v}} \\ &= \sqrt{\frac{1.055 \times 10^{-34}\ J \cdot s}{2\pi \cdot 1.139 \times 10^{-26}\ \text{kg} \cdot 1.153 \times 10^{11}\ \text{Hz}}} \\ &= 1.131 \times 10^{-10}\text{m} \\ &= 1.131\ \mathring{A}\end{aligned}$$

Advanced: Non-Rigid Rotors

Centrifugal stretching of the bond as J increases causes the decrease in the spacing between the lines in an observed spectrum (Table 5.9.1). This decrease shows that the molecule is not really a rigid rotor. As the rotational angular momentum increases with increasing J, the bond stretches. This stretching increases the moment of inertia and decreases the rotational constant (Figure 5.9.5).



Figure 5.9.5 : Two atoms connected by a vibrating bond. There is a rotation around the common center of mass, and oscillation in bond distance. (CC SA-BY 3.0; www.cleonis.nl).

The effect of centrifugal stretching is smallest at low J values, so a good estimate for B can be obtained from the $J = 0$ to $J = 1$ transition. From B, a value for the bond length of the molecule can be obtained since the moment of inertia that appears in the definition of B (Equation 5.9.16) is the reduced mass times the bond length squared. When the centrifugal stretching is taken into account quantitatively, the development of which is beyond the scope of the discussion here, a very accurate and precise

value for B can be obtained from the observed transition frequencies because of their high precision. Rotational transition frequencies are routinely reported to 8 and 9 significant figures.

As we have just seen, quantum theory successfully predicts the line spacing in a rotational spectrum. An additional feature of the spectrum is the line intensities. The lines in a rotational spectrum do not all have the same intensity, as can be seen in Figure 5.9.3 and Table 5.9.1. This is related to the populations of the initial and final states. This aspect of spectroscopy will be discussed in more detail in the following chapters

---

5.9: The Rigid Rotator is a Model for a Rotating Diatomic Molecule is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.9: The Rigid Rotator is a Model for a Rotating Diatomic Molecule** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

5.E: The Harmonic Oscillator and the Rigid Rotor (Exercises)

Solutions to select questions can be found online.

5.7

Calculate the reduced mass of HCl molecule given that the mass of H atom is 1.0078 amu and the mass of Cl atom is 34.9688 amu. Note that 1 amu = 1.660565*10^{-27} kg.

5.8

Calculate the reduced mass for the Br_2, Cl_2, and I_2 diatomics.

5.14

$^{79}Br^{79}Br$ has a force constant of $240\ N \cdot m^{-1}$. Given this information:

a. Calculate the fundamental vibrational frequency and
b. Calculate the $^{79}Br^{79}Br$ zero point energy.

5.19

Prove that the second derivative of an even function is even and odd function is odd.

5.27

The Harmonic oscillator Hamiltonian obeys the reflective property:

$$\hat{H}(x) = \hat{H}(-x)$$

What does this say about the nature of the harmonic oscillator wave function?

5.28

If $\langle x \rangle$ is an odd function, what does that say about p_x?

Hint: use

$$\frac{d\langle p_x \rangle}{dt} = \left\langle \frac{-dV}{dx} \right\rangle$$

also known as **Ehrenfest's Theorem**, where V is the potential of a one dimensional harmonic oscillator.

Hence, $\langle p_x \rangle$ does not depend on time.

5.32

Convert ∇^2 from Cartesian coordinates to cylindrical coordinates.

5.37

Find the magnitude of angular momentum and the z component of angular momentum for electrons in a hydrogen-like species with

a. quantum numbers $n = 1, l = 0, m = 0$; and
b. $n = 2, l = 0, m = 0$.

Compare your answers and explain your results.

5.38

Apply the angular momentum operator in the x direction to the following functions ($Y(\theta, \phi)$).

a. $\frac{5\pi}{4} + 7\exp(\pi^2)$
b. $3\pi \sin(\theta)$
c. $\frac{3}{2}\cos(\theta)\exp(i\phi)$

5.41

Use the fact that $\hat{x}$ and $\hat{p}$ are Hermitian in the number operator

$$\hat{a}_- = \frac{1}{\sqrt{2}}(\hat{x} + i\hat{p})$$

$$\hat{a}_+ = \frac{1}{\sqrt{2}}(\hat{x} - i\hat{p})$$

and

$$\hat{H} = \frac{\hbar w}{2}(\hat{a}_-\hat{a}_+ + \hat{a}_+\hat{a}_-)$$

Show that

$$\int \psi_v^* \hat{v} \psi dx \geq 0$$

5.43

Determine the unnormalized wave function $\psi_\circ(x)$ given that $\hat{a}_- = 2^{-1/2}(\hat{x} + i\hat{p})$ and that $\hat{a_-}\psi_\circ = 0$ Then find the unnormalized wave function for $\psi_1(x)$ using $\hat{a}_+$.

5.46

Find the reduced mass of an electron in a Tritium atom. Set the mass of the Tritium to be $5.008267 \times 10^{-27}\ kg$. Then find the value of the Rydberg constant for the Tritium atom.

5.46

The mass of a deuterium atom is $3.343586 \times 10^{-27}\ kg$. First calculate the reduce mass of the deuterium atom. Then using the reduced mass calculated find the Rydberg constant for a deuterium atom.

5.47

What is the ratio of the frequency of spectral lines of C-14 that has been ionized 5 times and C-12 that has been ionized 5 times?

5.47

Calculate the Rydberg constant for a deuterium atom and atomic hydrogen given the reduced mass of a deuterium atom is $9.106909 \times 10^{-31} kg$ and the reduced mass of hydrogen is $9.104431 \times 10^{-31} kg$. Compare both of these answers with the experimental result ($109677.6 cm^{-1}$). Then determine the ratio of the frequencies of the lines in the spectra of atomic hydrogen and atomic deuterium.

5.46

Find the reduced mass of HCl where the mass of hydrogen in 1 amu and the mass of chloride is 35 amu.

5.E: The Harmonic Oscillator and the Rigid Rotor (Exercises) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

6: The Hydrogen Atom

The solution of the Schrödinger equation (wave equation) for the hydrogen atom uses the fact that the Coulomb potential produced by the nucleus is isotropic (it is radially symmetric in space and only depends on the distance to the nucleus). Although the resulting energy eigenfunctions (the orbitals) are not necessarily isotropic themselves, their dependence on the angular coordinates follows completely generally from this isotropy of the underlying potential: the eigenstates of the Hamiltonian (that is, the energy eigenstates) can be chosen as simultaneous eigenstates of the angular momentum operator. This corresponds to the fact that angular momentum is conserved in the orbital motion of the electron around the nucleus. Therefore, the energy eigenstates may be classified by two angular momentum quantum numbers, ℓ and m (both are integers). The angular momentum quantum number $\ell = 0, 1, 2, \ldots$ determines the magnitude of the angular momentum. The magnetic quantum number $m = -\ell, \ldots, +\ell$ determines the projection of the angular momentum on the (arbitrarily chosen) z-axis.

6.1: The Schrödinger Equation for the Hydrogen Atom Can Be Solved Exactly

6.2: The Wavefunctions of a Rigid Rotator are Called Spherical Harmonics

6.3: The Three Components of Angular Momentum Cannot be Measured Simultaneously with Arbitrary Precision

6.4: Hydrogen Atomic Orbitals Depend upon Three Quantum Numbers

6.5: s Orbitals are Spherically Symmetric

6.6: Orbital Angular Momentum and the p-orbitals

6.7: The Helium Atom Cannot Be Solved Exactly

6.E: The Hydrogen Atom (Exercises)

Thumbnail: Depiction of a hydrogen atom with size of central proton shown, and the atomic diameter shown as about twice the Bohr model radius (image not to scale). Image used with permission (Public Domain; Bensaccount).

6: The Hydrogen Atom is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6.1: The Schrödinger Equation for the Hydrogen Atom Can Be Solved Exactly

The hydrogen atom, consisting of an electron and a proton, is a two-particle system, and the internal motion of two particles around their center of mass is equivalent to the motion of a single particle with a reduced mass. This reduced particle is located at r, where r is the vector specifying the position of the electron relative to the position of the proton. The length of r is the distance between the proton and the electron, and the direction of r and the direction of r is given by the orientation of the vector pointing from the proton to the electron. Since the proton is much more massive than the electron, we will assume throughout this chapter that the reduced mass equals the electron mass and the proton is located at the center of mass.

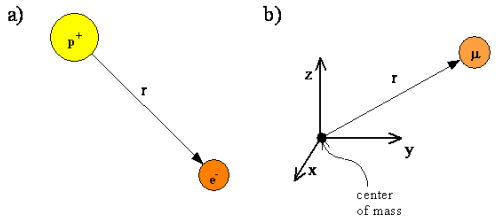


Figure 6.1.1 : (a) The proton (p^+) and electron (e^-) of the hydrogen atom. (b) Equivalent reduced particle with reduced mass μ at distance r from center of mass.

? Exercise 6.1.1

a. Assuming the Bohr radius gives the distance between the proton and electron, calculate the distance of the proton from the center of mass, and calculate the distance of the electron from the center of mass.
b. Calculate the reduced mass of the electron-proton system.
c. In view of your calculations in (a) and (b), comment on the validity of a model in which the proton is located at the center of mass and the reduced mass equals the electron mass.

Since the internal motion of any two-particle system can be represented by the motion of a single particle with a reduced mass, the description of the hydrogen atom has much in common with the description of a diatomic molecule discussed previously. The time-independent Schrödinger Equation for the hydrogen atom

$$\hat{H}(r,\theta,\varphi)\psi(r,\theta,\varphi) = E\psi(r,\theta,\varphi) \tag{6.1.1}$$

employs the same kinetic energy operator, $\hat{T}$, written in spherical coordinates. For the hydrogen atom, however, the distance, r, between the two particles can vary, unlike the diatomic molecule where the bond length was fixed and the rigid rotor model was applicable. The hydrogen atom Hamiltonian also contains a potential energy term, $\hat{V}$, to describe the attraction between the proton and the electron. This term is the Coulomb potential energy,

$$\hat{V}(r) = -\frac{e^2}{4\pi\epsilon_0 r} \tag{6.1.2}$$

where r is the distance between the electron and the proton. The Coulomb potential energy depends inversely on the distance between the electron and the nucleus and does not depend on any angles. Such a potential is called a **central potential**.

The full expression for $\hat{H}$ in spherical coordinates is

$$\hat{H}(r,\theta,\varphi) = -\frac{\hbar^2}{2\mu r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] - \frac{e^2}{4\pi\epsilon_0 r} \tag{6.1.3}$$

The contributions from rotational and radial components of the motion become clearer if we write out the complete Schrödinger equation,

$$\left\{-\frac{\hbar^2}{2\mu r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] - \frac{e^2}{4\pi\epsilon_0 r}\right\}\psi(r,\theta,\varphi) = E\psi(r,\theta,\varphi) \tag{6.1.4}$$

multiply both sides of Equation 6.1.4 by $2\mu r^2$, and rearrange to obtain

$$\hbar^2 \frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\psi(r,\theta,\varphi)\right) + 2\mu r^2\left[E + \frac{e^2}{4\pi\epsilon_0 r}\right]\psi(r,\theta,\varphi) =$$

$$-\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]\psi(r,\theta,\varphi) \tag{6.1.5}$$

Manipulating the Schrödinger equation in this way helps us recognize the square of the angular momentum operator in Equation 6.1.5. The square of the angular momentum operator in Equation 6.1.6.

$$\hat{M}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] \tag{6.1.6}$$

Substituting Equation 6.1.6 into Equation 6.1.5 produces

$$\hbar^2 \frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\psi(r,\theta,\varphi)\right) + 2\mu r^2[E - \hat{V}]\psi(r,\theta,\varphi) = \hat{M}^2\psi(r,\theta,\varphi) \tag{6.1.7}$$

? Exercise 6.1.2

Show the algebraic steps going from Equation 6.1.4 to 6.1.5 and finally to 6.1.7. Justify the statement that the rotational and radial motion are separated in Equation 6.1.7.

Since the angular momentum operator does not involve the radial variable, r, we can separate variables in Equation 6.1.7 by using a product wavefunction, as we did previously for rigid rotors. We know that the eigenfunctions of the angular momentum operator are the Spherical Harmonic functions (Table M4), $Y(\theta,\varphi)$, so a good choice for a product function is

$$\psi(r,\theta,\varphi) = R(r)Y(\theta,\varphi) \tag{6.1.8}$$

> *The Spherical Harmonic functions provide information about where the electron is around the proton, and the radial function $R(r)$ describes how far the electron is away from the proton.*

To separate variables, substitute the product function, Equation 6.1.8 into Equation 6.1.7, evaluate partial derivatives, divide each side by R(r) $Y(\theta,\varphi)$, and set each side of that resulting equation equal to a constant λ.

$$\frac{\hbar^2}{R(r)}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}R(r) + \frac{2\mu r^2}{R(r)}[E - V]R(r) = \lambda \tag{6.1.9}$$

$$\frac{1}{Y(\theta,\varphi)}\hat{M}^2 Y(\theta,\varphi) = \lambda \tag{6.1.10}$$

Equations 6.1.9 and 6.1.10 represent the radial differential equation and the angular differential equation, respectively. As we describe below, they are solved separately to give the $Y(\theta,\varphi)$ angular functions and the R(r) radial functions.

? Exercise 6.1.3

Complete the steps leading from Equations 6.1.7 to 6.1.9 and 6.1.10.

Rearranging Equation 6.1.10 yields

$$\hat{M}^2 Y_l^{m_l}(\theta,\varphi) = \lambda Y_l^{m_l}(\theta,\varphi) \tag{6.1.11}$$

where we have added the indices l and m_l to identify a particular spherical harmonic function. Note that the notation has changed from that used with the Rigid Rotor; it is customary to use J and m_J to represent the angular momentum quantum numbers for rotational states, but for electronic states, it is customary to use l and m_l to represent the same thing. Furthermore, the electronic angular momentum is designated by L and the corresponding operator is called $\hat{L}$. In complete electronic notation, Equation 6.1.11 is

$$\hat{L}^2 Y_l^{m_l}(\theta,\varphi) = \lambda Y_l^{m_l}(\theta,\varphi) \tag{6.1.12}$$

Equation 6.1.12 says that $Y_l^{m_l}(\theta,\varphi)$ must be an eigenfunction of the angular momentum operator $\hat{L}^2$ with eigenvalue λ. We know from the discussion of the Rigid Rotor that the eigenvalue λ is $J(J+1)\hbar^2$, or in electronic notation, $l(l+1)\hbar^2$. Consequently, Equation 6.1.12 becomes

$$\hat{L}^2 Y_l^{m_l}(\theta,\varphi) = l(l+1)\hbar^2 Y_l^{m_l}(\theta,\varphi) \tag{6.1.13}$$

Using this value for λ and rearranging Equation 6.1.9, we obtain

$$-\frac{\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}R(r) + \left[\frac{l(l+1)\hbar^2}{2\mu r^2} + V(r) - E\right]R(r) = 0 \tag{6.1.14}$$

? Exercise 6.1.4

Write the steps leading from Equation 6.1.9 to Equation 6.1.14.

The details for solving Equation 6.1.14 are provided elsewhere, but the procedure and consequences are similar to previously examined cases. As for the harmonic oscillator, an asymptotic solution (valid at large r) is found, and then the complete solutions are written as products of the asymptotic solution and polynomials arising from sequential truncations of a power series expansion.

Contributors and Attributions

- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")

6.1: The Schrödinger Equation for the Hydrogen Atom Can Be Solved Exactly is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.1: The Schrodinger Equation for the Hydrogen Atom Can Be Solved Exactly** is licensed CC BY-NC-SA 4.0.

6.2: The Wavefunctions of a Rigid Rotator are Called Spherical Harmonics

The solutions to the hydrogen atom Schrödinger equation are functions that are products of a spherical harmonic function and a radial function:

$$\psi_{n,l,m_l}(r,\theta,\phi) = R_{n,l}(r)Y_l^{m_l}(\theta,\phi) \tag{6.2.1}$$

The wavefunctions for the hydrogen atom depend upon the three variables r, θ, and ϕ and the three quantum numbers n, l, and m_l. The variables give the position of the electron relative to the nucleus in spherical coordinates. The absolute square of the wavefunction, $|\psi(r,\theta,\phi)|^2$, evaluated at r, θ, and ϕ gives the probability density of finding the electron inside a differential volume $d\tau$, centered at the position specified by r, θ, and ϕ.

? Exercise 6.2.1

What is the value of the integral (in braket notation)

$$\langle\psi(r,\theta,\phi)|\psi(r,\theta,\phi)\rangle \ ?$$

or expanded in integral notation (in Cartesian coordinates)

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}|\psi(x,y,z)|^2\,dx\,dy\,dz$$

or expanded in integral notation (in spherical coordinates)

$$\int_{0}^{+\infty}\int_{0}^{\pi}\int_{0}^{+2\pi}|\psi(r,\theta,\phi)|^2 r^2\sin\theta\,d\theta\,d\phi\,dr$$

The quantum numbers have names:

- n is called the principal quantum number,
- l is called the angular momentum quantum number, and
- m_l is called the magnetic quantum number because (as we will see, the energy in a magnetic field depends upon m_l).

Often l is called the azimuthal quantum number because it is a consequence of the θ-equation, which involves the azimuthal angle Θ, referring to the angle to the zenith. These quantum numbers have specific values that are dictated by the physical constraints or boundary conditions imposed upon the Schrödinger equation: n must be an integer greater than 0, l can have the values 0 to n-1, and m_l can have $2l+1$ values ranging from $-l$ to $+l$ in unit or integer steps. The total number of orbitals with a particular value of n is n^2 - i.e., this is the **degeneracy** of the system.

Wavefunction Nomenclature

The values of the quantum number l usually are coded by a letter: s means 0, p means 1, d means 2, f means 3; the next codes continue alphabetically (e.g., g means $l = 4$). The quantum numbers specify the quantization of physical quantities. The discrete energies of different states of the hydrogen atom are given by n, the magnitude of the angular momentum is given by l, and one component of the angular momentum (usually chosen by chemists to be the z-component) is given by m_l.

The Φ function is found to have the quantum number m, where

$$\Phi_m(\phi) = A_m e^{im\phi}$$

and A_m is the normalization constant and $m = 0, \pm1, \pm2... \pm\infty$. The Θ function was solved and is known as *Legendre polynomials*, which have quantum numbers m and ℓ. When Θ and Φ are multiplied together, the product is known as **spherical harmonics** with labeling $Y_J^m(\theta,\phi)$.

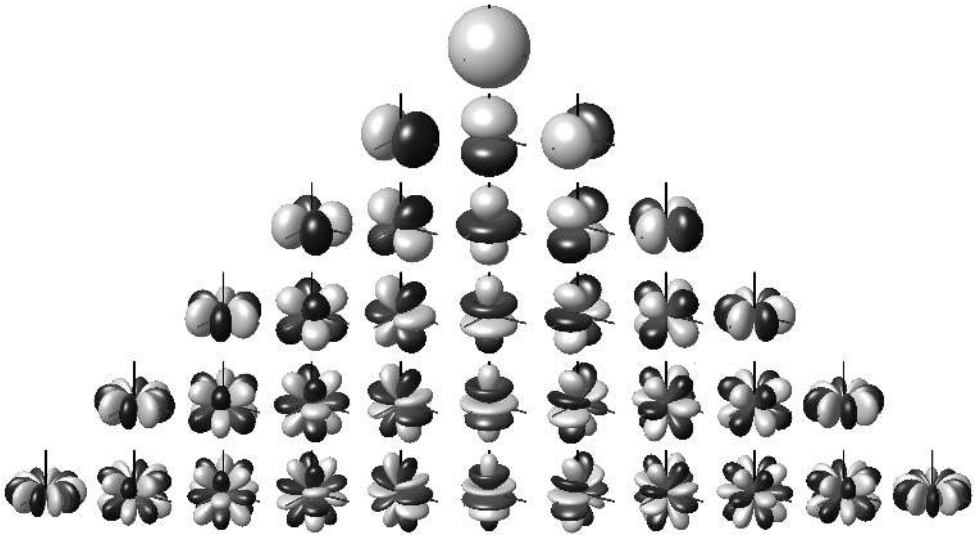


Figure 6.2.1: Spherical Harmonics as commonly displayed, sorted by increasing energies and aligned for symmetry.

Figure 6.2.1 shows the spherical harmonics Y_J^M, which are solutions of the angular Schrödinger equation. These are explicitly written in Table 6.2.1 . Notice that these functions are complex in nature.

Table 6.2.1 : spherical harmonics Y_J^M

m_J	J	$\Theta_J^{m_J}(\theta)$	$\Phi(\varphi)$	$Y_J^{m_J}(\theta,\varphi)$
0	0	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2\pi}}$	$\frac{1}{\sqrt{4\pi}}$
0	1	$\sqrt{\frac{3}{2}}\cos\theta$	$\frac{1}{\sqrt{2\pi}}$	$\sqrt{\frac{3}{4\pi}}\cos\theta$
1	1	$\sqrt{\frac{3}{4}}\sin\theta$	$\frac{1}{\sqrt{2\pi}}e^{i\varphi}$	$\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\varphi}$
-1	1	$\sqrt{\frac{3}{4}}\sin\theta$	$\frac{1}{\sqrt{2\pi}}e^{-i\varphi}$	$\sqrt{\frac{3}{8\pi}}\sin\theta e^{-i\varphi}$
0	2	$\sqrt{\frac{5}{8}}(3\cos^2\theta-1)$	$\frac{1}{\sqrt{2\pi}}$	$\sqrt{\frac{5}{16\pi}}(3\cos^2\theta-1)$
1	2	$\sqrt{\frac{15}{4}}\sin\theta\cos\theta$	$\frac{1}{\sqrt{2\pi}}e^{i\varphi}$	$\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{i\varphi}$
-1	2	$\sqrt{\frac{15}{4}}\sin\theta\cos\theta$	$\frac{1}{\sqrt{2\pi}}e^{-i\varphi}$	$\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{-i\varphi}$
2	2	$\sqrt{\frac{15}{16}}\sin^2\theta$	$\frac{1}{\sqrt{2\pi}}e^{2i\varphi}$	$\sqrt{\frac{15}{32\pi}}\sin^2\theta e^{2i\varphi}$
-2	2	$\sqrt{\frac{15}{16}}\sin^2\theta$	$\frac{1}{\sqrt{2\pi}}e^{-2i\varphi}$	$\sqrt{\frac{15}{32\pi}}\sin^2\theta e^{-2i\varphi}$

? Exercise 6.2.2

Consider several values for n, and show that the number of orbitals for each n is n^2.

? Exercise 6.2.3

Construct a table summarizing the allowed values for the quantum numbers n, l, and m_l. for energy levels 1 through 7 of hydrogen.

? Exercise 6.2.4

The notation 3d specifies the quantum numbers for an electron in the hydrogen atom.

a. What are the values for n and l?

b. What are the values for the energy and angular momentum?
c. What are the possible values for the magnetic quantum number?
d. What are the possible orientations for the angular momentum vector?

Atomic Orbitals

The hydrogen atom wavefunctions, $\psi(r,\theta,\phi)$, are called **atomic orbitals**. An atomic orbital is a function that describes one electron in an atom. The wavefunction with $n = 1$, $l = 0$, and $m_l = 0$ is called the 1s orbital, and an electron that is described by this function is said to be "in" the 1s orbital, i.e. have a 1s orbital state. The constraints on n, ℓ, and m_ℓ that are imposed during the solution of the hydrogen atom Schrödinger equation explain why there is a single 1s orbital, why there are three 2p orbitals, five 3d orbitals, etc. We will see when we consider multi-electron atoms, these constraints explain the features of the Periodic Table. In other words, the Periodic Table is a manifestation of the Schrödinger model and the physical constraints imposed to obtain the solutions to the Schrödinger equation for the hydrogen atom.

The first few solutions of the Schrödinger equation for the H atom are:

Table 6.2.1 : The solutions to the Schrödinger equation for one-electron atoms and ions. Z is the atomic number of the nucleus, and $\rho = \frac{Zr}{a_0}$, where a_0 is the Bohr radius and r is the radial variable.

	n	l	m_l	ψ_{n,l,m_l}
ψ_{100}	1	0	$m_l = 0$	$\frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-\frac{Zr}{a_0}}$
ψ_{200}	2	0	$m_l = 0$	$\frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}}\left(2-\frac{Zr}{a_0}\right) e^{-\frac{Zr}{2a_0}}$
ψ_{210}	2	1	$m_l = 0$	$\frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} re^{-\frac{Zr}{2a_0}}\cos\theta$
$\psi_{21\pm1}$	2	1	$m_l = \pm1$	$\frac{1}{\sqrt{64\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} re^{-\frac{Zr}{2a_0}}\sin\theta e^{\pm i\phi}$
ψ_{300}	3	0	$m_l = 0$	$\frac{1}{81\sqrt{3\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}}\left(27-18\frac{Zr}{a_0}+2\frac{Z^2r^2}{a_0^2}\right.$
ψ_{310}	3	1	$m_l = 0$	$\frac{\sqrt{2}}{81\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} r\left(6-\frac{Zr}{a_0}\right) e^{-\frac{Zr}{3a_0}}\cos$
$\psi_{31\pm1}$	3	1	$m_l = \pm1$	$\frac{1}{81\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} r\left(6-\frac{Zr}{a_0}\right) e^{-\frac{Zr}{3a_0}}\sin$
ψ_{320}	3	2	$m_l = 0$	$\frac{1}{81\sqrt{6\pi}}\left(\frac{Z}{a_0}\right)^{\frac{7}{2}} r^2 e^{-\frac{Zr}{3a_0}}\left(3\cos^2\theta-1\right.$
$\psi_{32\pm1}$	3	2	$m_l = \pm1$	$\frac{1}{81\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{7}{2}} r^2 e^{-\frac{Zr}{3a_0}}\sin\theta\cos\theta e^{\pm i\phi}$
$\psi_{32\pm2}$	3	2	$m_l = \pm2$	$\frac{1}{162\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{7}{2}} r^2 e^{-\frac{Zr}{3a_0}}\sin^2\theta e^{\pm 2i\phi}$

We can take superposition states of the solutions to obtain real wavefunctions that do not have any imaginary components:

Table 6.2.2 : The real solutions to the Schrödinger equation for one-electron atoms and ions. Z is the atomic number of the nucleus, and $\rho = \frac{Zr}{a_0}$, where a_0 is the Bohr radius and r is the radial variable.

	n	l	m_l	$\psi_{n,l}$
ψ_{1s}	1	0	0	$\frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-\frac{Zr}{a_0}}$
ψ_{2s}	2	0	0	$\frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}}\left(2-\frac{Zr}{a_0}\right)e^{-\frac{Zr}{2a_0}}$
ψ_{2p_z}	2	1	0	$\frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} re^{-\frac{Zr}{2a_0}}\cos\theta$
ψ_{2p_x}	2	1	±1	$\frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} re^{-\frac{Zr}{2a_0}}\sin\theta\cos\phi$
ψ_{2p_y}	2	1	±1	$\frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} re^{-\frac{Zr}{2a_0}}\sin\theta\sin\phi$
ψ_{3s}	3	0	0	$\frac{1}{81\sqrt{3\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}}\left(27-18\frac{Zr}{a_0}+2\frac{Z^2r^2}{a_0^2}\right.$
ψ_{3p_z}	3	1	0	$\frac{\sqrt{2}}{81\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} r\left(6-\frac{Zr}{a_0}\right)e^{-\frac{Zr}{3a_0}}\cos$
ψ_{3p_x}	3	1	±1	$\frac{\sqrt{2}}{81\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} r\left(6-\frac{Zr}{a_0}\right)e^{-\frac{Zr}{3a_0}}\sin$
ψ_{3p_y}	3	1	±1	$\frac{\sqrt{2}}{81\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} r\left(6-\frac{Zr}{a_0}\right)e^{-\frac{Zr}{3a_0}}\sin$
$\psi_{3d_{z^2}}$	3	2	0	$\frac{1}{81\sqrt{6\pi}}\left(\frac{Z}{a_0}\right)^{\frac{7}{2}} r^2e^{-\frac{Zr}{3a_0}}\left(3\cos^2\theta-1\right.$
$\psi_{3d_{xz}}$	3	2	±1	$\frac{\sqrt{2}}{81\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{7}{2}} r^2e^{-\frac{Zr}{3a_0}}\sin\theta\cos\theta\cos$
$\psi_{3d_{yz}}$	3	2	±1	$\frac{\sqrt{2}}{81\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{7}{2}} r^2e^{-\frac{Zr}{3a_0}}\sin\theta\cos\theta\sin$
$\psi_{3d_{x^2-y^2}}$	3	2	±2	$\frac{1}{81\sqrt{2\pi}}\left(\frac{Z}{a_0}\right)^{\frac{7}{2}} r^2e^{-\frac{Zr}{3a_0}}\sin^2\theta\cos 2\phi$
$\psi_{3d_{xy}}$	3	2	±2	$\frac{1}{81\sqrt{2\pi}}\left(\frac{Z}{a_0}\right)^{\frac{7}{2}} r^2e^{-\frac{Zr}{3a_0}}\sin^2\theta\sin 2\phi$

These are the wavefunctions that are most often depicted. Visualizing the variation of an electronic wavefunction with r, θ, and ϕ is important because the absolute square of the wavefunction depicts the charge distribution (electron probability density) in an atom or molecule. The charge distribution is central to chemistry because it is related to chemical reactivity. For example, an electron-deficient part of one molecule is attracted to an electron-rich region of another molecule, and such interactions play a major role in chemical interactions ranging from substitution and addition reactions to protein folding and the interaction of substrates with enzymes.

Visualizing wavefunctions and charge distributions is challenging because it requires examining the behavior of a function of three variables in three-dimensional space. This visualization is made easier by considering the radial and angular parts separately, but plotting the radial and angular parts separately do not reveal the shape of an orbital very well. The shape can be revealed better in a probability density plot. To make such a three-dimensional plot, divide space up into small volume elements, calculate $\psi^*\psi$ at the center of each volume element, and then shade, stipple or color that volume element in proportion to the magnitude of $\psi^*\psi$. Do not confuse such plots with polar plots, which look similar. Probability densities also can be represented by contour maps, as shown in Figure 6.2.2 .

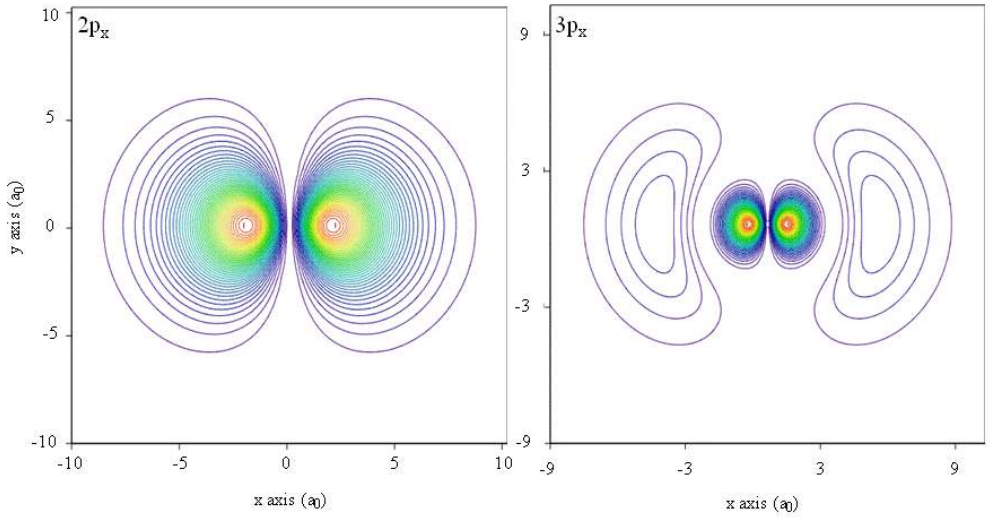


Figure 6.2.2: Contour plots in the x-y plane for the $2p_x$ and $3p_x$ orbitals of the hydrogen atom. The plots map lines of constant values of $R(r)^2$; red lines follow paths of high $R(r)^2$, blue for low $R(r)^2$. The angular function used to create the figure was a linear combination of two Spherical Harmonic functions.

Methods for separately examining the radial portions of atomic orbitals provide useful information about the distribution of charge density within the orbitals. Graphs of the radial functions, $R(r)$, for the 1s, 2s, and 2p orbitals plotted in Figure 6.2.3 .

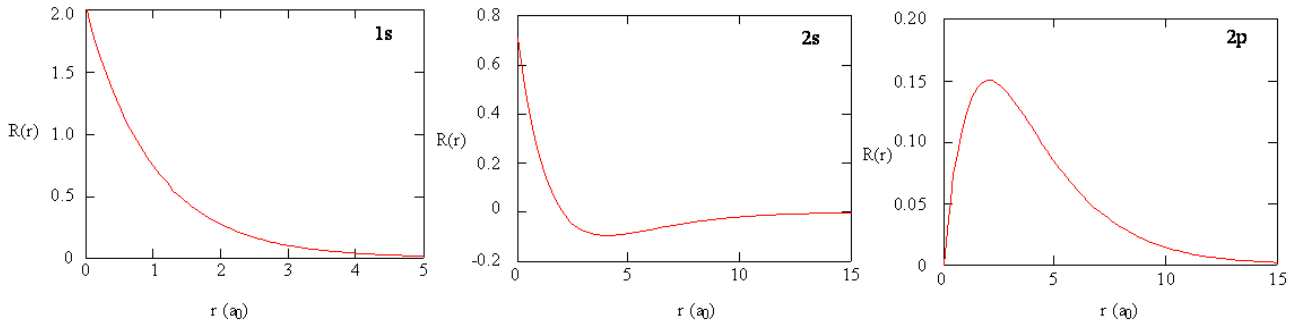


Figure 6.2.3: Radial function, $R(r)$, for the 1s, 2s, and 2p orbitals.

The 1s function in Figure 6.2.3 starts with a high positive value at the nucleus and exponentially decays to essentially zero after 5 Bohr radii. The high value at the nucleus may be surprising, but as we shall see later, the probability of finding an electron at the nucleus is vanishingly small.

Next notice how the radial function for the 2s orbital, Figure 6.2.3 , goes to zero and becomes negative. This behavior reveals the presence of a radial node in the function. A radial node occurs when the radial function equals zero other than at $r = 0$ or $r = \infty$. Nodes and limiting behaviors of atomic orbital functions are both useful in identifying which orbital is being described by which wavefunction. For example, all of the s functions have non-zero wavefunction values at $r = 0$, but p, d, f and all other functions go to zero at the origin. It is useful to remember that there are $n - 1 - l$ radial nodes in a wavefunction, which means that a 1s orbital has no radial nodes, a 2s has one radial node, and so on.

? Exercise 6.2.5

Examine the mathematical forms of the radial wavefunctions. What feature in the functions causes some of them to go to zero at the origin while the s functions do not go to zero at the origin?

? Exercise 6.2.6

What mathematical feature of each of the radial functions controls the number of radial nodes?

? Exercise 6.2.7 : Radial Node

At what value of r does the 2s radial node occur?

Exercise 6.2.8

Make a table that provides the energy, number of radial nodes, and the number of angular nodes and total number of nodes for each function with $n = 1$, $n = 2$, and $n = 3$. Identify the relationship between the energy and the number of nodes. Identify the relationship between the number of radial nodes and the number of angular nodes.

The quantity $R(r)^*R(r)$ gives the radial probability density; i.e., the probability density for the electron to be at a point located the distance r from the proton. Radial probability densities for three atomic orbitals are plotted in Figure 6.2.4 .

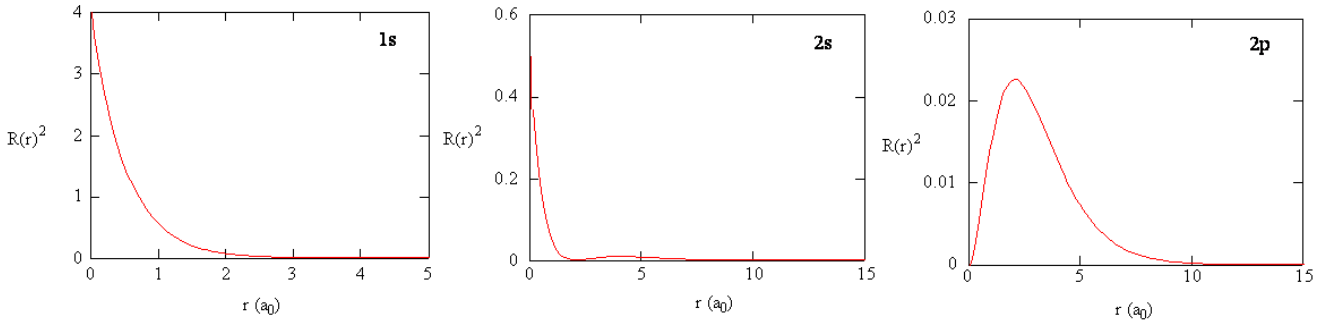


Figure 6.2.4: Radial probability densities for the 1s, 2s, and 2p orbitals.

Probabilities and Distribution Functions

When the **radial probability density** for every value of r is multiplied by the area of the spherical surface represented by that particular value of r, we get the radial distribution function. The radial distribution function gives the probability density for an electron to be found anywhere on the surface of a sphere located a distance r from the proton. Since the area of a spherical surface is $4\pi r^2$, the **radial distribution function** is given by

$$\underbrace{4\pi r^2 R(r)^* R(r)}_{\text{radial probability density}}$$

Radial distribution functions are shown in Figure 6.2.5 . At small values of r, the radial distribution function is low because the small surface area for small radii modulates the high value of the radial probability density function near the nucleus. As we increase r, the surface area associated with a given value of r increases, and the r^2 term causes the radial distribution function to increase even though the radial probability density is beginning to decrease. At large values of r, the exponential decay of the radial function outweighs the increase caused by the r^2 term and the radial distribution function decreases.

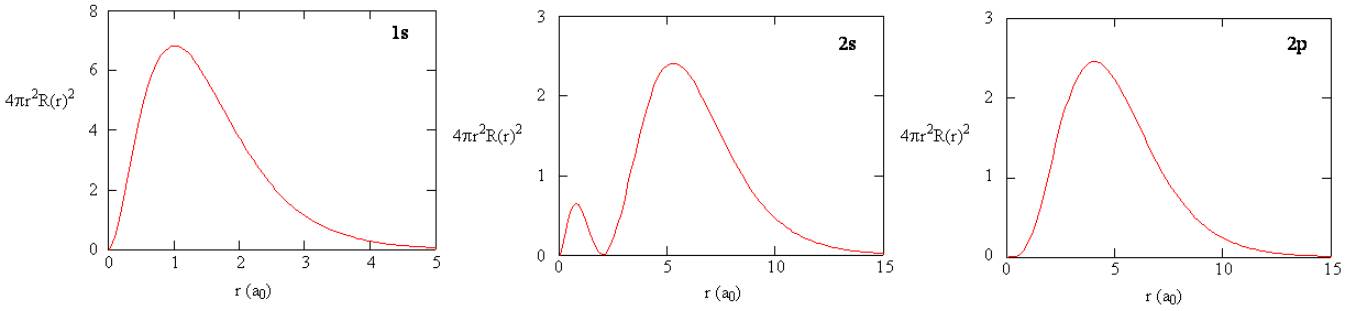


Figure 6.2.5: The radial distribution function for the 1s, 2s, and 2p orbitals.

Exercise 6.2.9

Write a quality comparison of the radial function and radial distribution function for the 2s orbital. See Figure 6.2.6 .

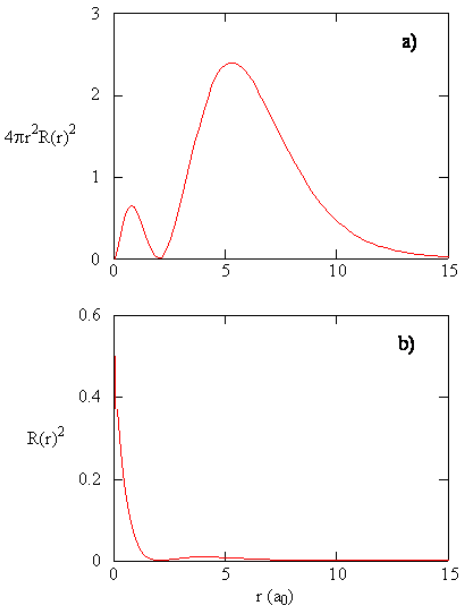


Figure 6.2.6: Comparison of a) the radial distribution function and b) the radial probability density for the 2s orbital.

6.2: The Wavefunctions of a Rigid Rotator are Called Spherical Harmonics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.2: The Wavefunctions of a Rigid Rotator are Called Spherical Harmonics** is licensed CC BY-NC-SA 4.0.
- **6.4: Hydrogen Atomic Orbitals Depend upon Three Quantum Numbers** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

6.3: The Three Components of Angular Momentum Cannot be Measured Simultaneously with Arbitrary Precision

Learning Objectives

- Understand how to measure the orbital angular momentum of an electron around a nucleus.
- Understand how the Heisenburgh Uncertainty Principle extends to orbital angular momenta.
- Manipulate the angular momenta cyclic permutations that allow two of the three projects to be simultaneous measured

Consider a particle described by the Cartesian coordinates $(x, y, z) \equiv \vec{r}$ and their conjugate momenta $(p_x, p_y, p_z) \equiv \vec{p}$. The classical definition of the *orbital angular momentum* of such a particle about the origin is (i.e., via the vector cross product):

$$\vec{L} = \vec{r} \times \vec{p}$$

which can be separated into projections into each of the primary axes :

Multiple \label

Extending this discussion to the quantum mechanics, we can assume that the operators $(\hat{L}_x, \hat{L}_y, \hat{L}_z) \equiv \vec{L}$ - that represent the components of orbital angular momentum in quantum mechanics - can be defined in an analogous manner to the corresponding components of classical angular momentum. In other words, we are going to assume that the above equations specify the angular momentum operators in terms of the position and linear momentum operators.

In Cartesian coordinates, the three operators for the orbital angular momentum components can be written as

$$\hat{L}_x = -\mathrm{i}\,\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \tag{6.3.1}$$

$$\hat{L}_y = -\mathrm{i}\,\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \tag{6.3.2}$$

$$\hat{L}_z = -\mathrm{i}\,\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \tag{6.3.3}$$

These can be transforming to operators in standard spherical polar coordinates,

Multiple \label

we obtain

Multiple \label

We can introduce a new operator $\hat{L}^2$:

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \tag{6.3.4}$$

$$= -\hbar^2 \left(\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \sin\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right) \tag{6.3.5}$$

The eigenvalue problem for $\hat{L}^2$ takes the form

$$\hat{L}^2 |\psi\rangle = \lambda\,\hbar^2 |\psi\rangle \tag{6.3.6}$$

where $\psi(r, \theta, \varphi)$ is the wavefunction, and λ is a number. Let us write

$$\psi(r, \theta, \varphi) = R(r)\, Y(\theta, \varphi) \tag{6.3.7}$$

By definition,

$$\boxed{L^2 Y_l^{m_l} = l\,(l+1)\,\hbar^2\, Y_l^{m_l}} \tag{6.3.8}$$

where l is an integer. This is an important conclusion that argues the angular momentum is quantized with the square of the magnitude of the angular momentum only capable of assume one of the discrete set of values (Equation 6.3.8). From this, the **amplitude** of angular momentum can be expressed

$$\boxed{|\vec{L}| = \sqrt{L^2} = \sqrt{l(l+1)}\hbar} \tag{6.3.9}$$

Warning

We often refer to a particle in a state with angular momentum quantum number l as having angular momentum l, rather than saying that it has angular momentum of $\sqrt{l(l+1)}\hbar$ magnitude, primarily since it is awkward to say quickly.

The properties of spherical harmonics that the z-component of the angular momentum (L_z) is also quantized and can only assume a one of a discrete set of values

$$L_z Y_l^{m_l} = m\hbar Y_l^{m_l} \tag{6.3.10}$$

where m_l is an integer lying in the range $-l \leq m_l \leq l$.

- l is sometimes called "azimuthal quantum number" or "orbital quantum number"
- m_l is sometimes called "magnetic quantum number"

Simultaneous Measurements

Note that observables associated with $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ can, in principle, be measured *simultaneously* with infinite precision, the corresponding operators must **commute**. Remember that the fundamental commutation relations satisfied by the position and linear momentum operators are:

Multiple \label

where i and j stand for either x, y, or z. Consider the commutator of the operators $\hat{L}_x$ and $\hat{L}_z$:

Multiple \label

The **cyclic permutations** of the above result yield the fundamental commutation relations satisfied by the components of an orbital angular momentum:

$$[\hat{L}_x, \hat{L}_y] = \mathrm{i}\hbar\hat{L}_z \tag{6.3.11}$$

$$[\hat{L}_y, \hat{L}_z] = \mathrm{i}\hbar\hat{L}_x \tag{6.3.12}$$

$$[\hat{L}_z, \hat{L}_x] = \mathrm{i}\hbar\hat{L}_y \tag{6.3.13}$$

The three commutation relations (Equations 6.3.11 - 6.3.13) are the foundation for the whole theory of angular momentum in quantum mechanics. Whenever we encounter three operators having these commutation relations, we know that the dynamical variables that they represent have identical properties to those of the components of an angular momentum (which we are about to derive). In fact, we shall assume that any three operators that satisfy the commutation relations (Equations 6.3.11 - 6.3.13) represent the components of some sort of angular momentum.

> *Any three operators that satisfy the cyclic commutation relations represent the components of some sort of angular momentum.*

Example 6.3.1 : Commutators

Show that the $\hat{L}^2$ and $\hat{L}_x$ operators commute.

Solution

We want to confirm that $[\hat{L}^2, \hat{L}_x] = 0$ that from Equation 6.3.4 this can be expanded

$$[\hat{L}^2, \hat{L}_x] = [\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \hat{L}_x]$$

from the properties of commutators, this can be expanded

$$[\hat{L}^2, \hat{L}_x] = [\hat{L}_x^2, \hat{L}_x] + [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x]$$

However,

$$[\hat{L}_x^2, \hat{L}_x] = \hat{L}_x^2\hat{L}_x - \hat{L}_x\hat{L}_x^2 = \hat{L}_x\hat{L}_x\hat{L}_x - \hat{L}_x\hat{L}_x\hat{L}_x = 0$$

So

$$\begin{aligned}[\hat{L}^2, \hat{L}_x] &= [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x] \\ &= \hat{L}_y^2\hat{L}_x - \hat{L}_x\hat{L}_y^2 + \hat{L}_z^2\hat{L}_x - \hat{L}_x\hat{L}_z^2 \\ &= \hat{L}_y\hat{L}_y\hat{L}_x - \hat{L}_x\hat{L}_y\hat{L}_y + \hat{L}_z\hat{L}_z\hat{L}_x - \hat{L}_x\hat{L}_z\hat{L}_z\end{aligned}$$

Lets look at some related forms which can be used to simplify the above expression. The first two terms can and final two terms can be rewritten as different commutators

$$\begin{aligned}[\hat{L}_y, \hat{L}_x] &= \hat{L}_y + \hat{L}_y[\hat{L}_y, \hat{L}_x] \\ &= (\hat{L}_y\hat{L}_x - \hat{L}_x\hat{L}_y)\hat{L}_y + \hat{L}_y(\hat{L}_y\hat{L}_x - \hat{L}_x\hat{L}_y) \\ &= \cancel{\hat{L}_y\hat{L}_x\hat{L}_y} - \hat{L}_x\hat{L}_y\hat{L}_y + \hat{L}_y\hat{L}_y\hat{L}_x - \cancel{\hat{L}_y\hat{L}_x\hat{L}_y}\end{aligned}$$

The first & fourth terms cancel, giving

$$[\hat{L}_y, \hat{L}_x]\hat{L}_y + \hat{L}_y[\hat{L}_y, \hat{L}_x] = \hat{L}_y\hat{L}_y\hat{L}_x - \hat{L}_x\hat{L}_y\hat{L}_y$$

Similarly,

$$[Lz, \hat{L}_x]\hat{L}_z + \hat{L}_z[\hat{L}_z, \hat{L}_x] = \hat{L}_z\hat{L}_z\hat{L}_x - \hat{L}_x\hat{L}_z\hat{L}_z$$

So,

$$\begin{aligned}[\hat{L}^2, \hat{L}_x] &= [\hat{L}_y, \hat{L}_x]\hat{L}_y + \hat{L}_y[\hat{L}_y, \hat{L}_x] + [\hat{L}_z, \hat{L}_x]\hat{L}_z + \hat{L}_z[\hat{L}_z, \hat{L}_x] \\ &= -i\hbar\hat{L}_z\hat{L}_y - i\hbar\hat{L}_y\hat{L}_z + i\hbar\hat{L}_y\hat{L}_z + ih\hat{L}_z\hat{L}_y \\ &= 0\end{aligned}$$

One can also show similarly that

$$[\hat{L}^2, \hat{L}_y] = [\hat{L}^2, \hat{L}_z] = 0$$

Example 6.3.1 shows that while L_z can be known with certainty, L_x and L_y would **unknown**. This means that every vector with the appropriate length and z-component can drawn to represent $\vec{L}$, which forms a cone (Figure 6.3.1). 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 while it cannot be defined for a single system (since the components of L do not commute with each other).

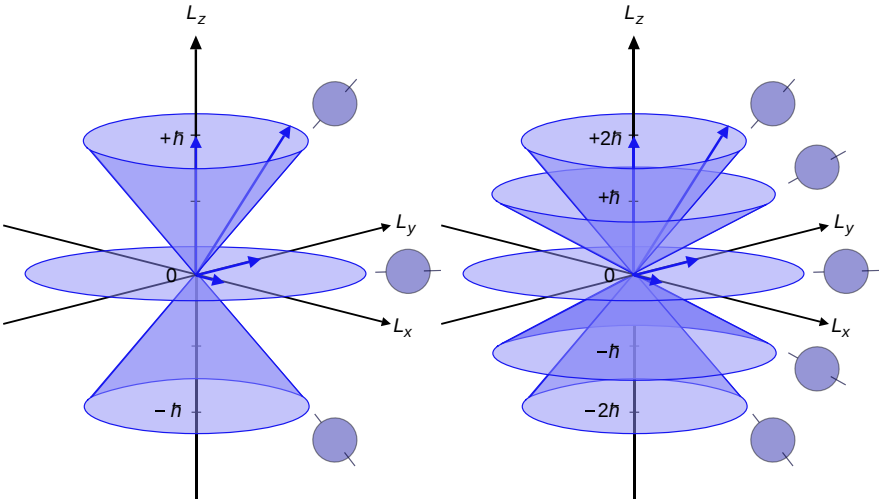


Figure 6.3.1 : Illustration of the vector model of orbital angular momentum. (left) A set of states with quantum numbers $l = 1$, and $m_l = -1, 0, +1,$. (right) A set of states with quantum numbers $l = 2$, and $m_l = -2, -1, 0, +1, +2$. (Public Domain; Maschen via Wikipedia)

The Meaning of Commutation of Two Operators

The mathematics of commutation relations is relatively straightforward, but what does it physically mean for an observable (Hermitian operator) to commute with another observable (Hermitian operator) in quantum mechanics?

If two operators $\hat{A}$ and $\hat{B}$ commute with each other then

$$\hat{A}\hat{B} - \hat{B}\hat{A} = 0,$$

which can be rearranged to

$$\hat{A}\hat{B} = \hat{B}\hat{A}.$$

This is not a trivial statement and many operations do not commute and hence the end-result depends on how you have ordered the operations.

If you recall that operators act on quantum mechanical states and give you a new state in return, then this means that with $\hat{A}$ and $\hat{B}$ commuting, the state you obtain from letting first $\hat{A}$ act and then $\hat{B}$ act on some initial state is the same as if you let first $\hat{B}$ and then $\hat{A}$ act on that state, i.e.,

$$\hat{A}\hat{B}|\psi\rangle = \hat{B}\hat{A}|\psi\rangle.$$

Recall that when you perform a quantum mechanical measurement, you will always measure an eigenvalue of your operator, and after the measurement your state is left in the corresponding eigenstate. The eigenstates to the operator are precisely those states for which there is **no uncertainty** in the measurement: You will **always** measure the eigenvalue.

Therefore, $\hat{B}|a\rangle$ must be an eigenfunction of $\hat{A}$ with eigenvalue a just like $|a\rangle$ itself is. That is essentially saying that $|a\rangle$ is an eigenfunction of $\hat{B}$.

A key example of this is since $\hat{L}^2$ and $\hat{L}_x$ commute (Example 6.3.1) then both operators share the same eigenstates. Hence, we do no need to solve two eignevalue problems:

$$\hat{L}^2|\psi\rangle = \lambda|\psi\rangle$$

and

$$\hat{L}_x|\psi\rangle = \beta|\psi\rangle$$

If we solve one, we then know the eigenvalues ($|\psi\rangle$) for the other!

What does it mean when some observable $\hat{A}$ commutes with the Hamiltonian $\hat{H}$? First, we get all the result from above: There is a simultaneous eigenbasis of the energy-eigenstates and the eigenstates of $\hat{A}$. This can yield a tremendous simplification of the task of solving Schrödinger equations. For example, the Hamiltonian of the hydrogen atom commutes with $\hat{L}$, the angular momentum operator, and with $\hat{L}_z$, its z-component. This tells you that you can classify the eigenstates by an angular- and magnetic quantum number l and m.

Summary

In the quantum world, angular momentum is quantized. The square of the magnitude of the angular momentum (determined by the eigenvalues of the $\hat{L}^2$ operator) can only assume one of the discrete set of values

$$L^2 = l(l+1)\hbar^2$$

or the magnitude of the angular momentum

$$L = \sqrt{l(l+1)}\hbar$$

with $l = 0, 1, 2, ...$

The z-component of the angular momentum (i.e., projection of L onto the z-axis) is also quantized with

$$L_z = m_l\hbar$$

with $m_l = -l, 0-1, ..., 0, ... + l + 1, l$ for a given value of l. Hence, l and m_l are the *angular momentum quantum number* and the *magnetic quantum number*, respectively.

Contributors and Attributions

- Richard Fitzpatrick (Professor of Physics, The University of Texas at Austin)

6.3: The Three Components of Angular Momentum Cannot be Measured Simultaneously with Arbitrary Precision is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.3: The Three Components of Angular Momentum Cannot be Measured Simultaneously with Arbitrary Precision** is licensed CC BY-NC-SA 4.0.

6.4: Hydrogen Atomic Orbitals Depend upon Three Quantum Numbers

Learning Objectives

- Recognize how the hydrogen atomic orbitals vary as a function of the three primary quantum numbers

The solutions to the hydrogen atom Schrödinger equation discussed previously are functions that are products of a **spherical harmonic function** and a **radial function**.

$$\psi_{n,l,m_l}(r,\theta,\varphi) = \underbrace{R_{n,l}(r)}_{radial}\underbrace{Y_l^{m_l}(\theta,\varphi)}_{angular} \tag{6.4.1}$$

The wavefunctions for the hydrogen atom depend upon the three variables r, θ, and φ and the three quantum numbers n, l, and m_l. The variables give the position of the electron relative to the proton in spherical coordinates. The absolute square of the wavefunction, $|\psi(r,\theta,\varphi)|^2$, evaluated at r, θ, and φ gives the probability density of finding the electron inside a differential volume $d\tau$, centered at the position specified by r, θ, and φ.

Exercise 6.4.1

Evaluate the following integrals

a. $\langle\psi(r,\theta,\varphi)|\psi(r,\theta,\varphi)\rangle$
b. $\langle\psi(r,\theta,\varphi)|\psi(r',\theta',\varphi')\rangle$

The quantum numbers n, l, and m_l have names:

- n is called the **principal quantum number**,
- l is called the **angular momentum quantum number**, and
- m_l is called the **magnetic quantum number** because the energy in a magnetic field depends upon m_l.

Often l is called the azimuthal quantum number because it is a consequence of the θ-equation, which involves the azimuthal angle Θ, referring to the angle to the zenith.

The three quantum numbers have specific values that are dictated by the physical constraints or boundary conditions imposed upon the Schrödinger equation: n must be an integer greater than 0, l can have the values 0 to n-1, and m_l can have $2l+1$ values ranging from $-l$ to $+l$ in unit or integer steps. The values of the quantum number l usually are coded by a letter: s means 0, p means 1, d means 2, f means 3; the next codes continue alphabetically (e.g., g means $l = 4$). The quantum numbers specify the quantization of physical quantities. The discrete energies of different states of the hydrogen atom are given by n, the magnitude of the angular momentum is given by l, and one component of the angular momentum (usually chosen by chemists to be the z-component) is given by m_l. The total number of orbitals with a particular value of n is n^2.

Exercise 6.4.4

Consider several values for n, and show that the number of orbitals for each n is n^2.

Exercise 6.4.5

Construct a table summarizing the allowed values for the quantum numbers n, l, and m_l for energy levels 1 through 7 of hydrogen.

Exercise 6.4.6

The notation 3d specifies the quantum numbers for an electron in the hydrogen atom. What are the values for n and l ? What are the values for the energy and angular momentum? What are the possible values for the magnetic quantum number? What are the possible orientations for the angular momentum vector?

Radial Part of the Wavefunction

The asymptotic behavior (i.e., far away from the nucleus) to the radial part of the wavefunction is

$$R_{\text{asymptotic}}(r) \sim \exp\left(-\frac{r}{n}a_0\right) \tag{6.4.2}$$

where n will turn out to be a quantum number and a_0 is the Bohr radius (~52.9 pm). Note that this function decreases exponentially with distance, in a manner similar to the decaying exponential portion of the harmonic oscillator wavefunctions, but with a different distance dependence, r vs. r^2.

? Exercise 6.4.2

What happens to the magnitude of $R_{\text{asymptotic}}(r)$ as the distance r from the proton approaches infinity? Sketch a graph of the function, $R_{\text{asymptotic}}(r)$. Why might this behavior be expected for an electron in a hydrogen atom?

The polynomials produced by the truncation of the power series are related to the associated Laguerre polynomials, $L_{n,l}(r)$, where the set of c_i are constant coefficients.

$$L_{n,l}(r) = \sum_{r=0}^{n-l-1} c_i r^i \tag{6.4.3}$$

These polynomials are identified by two indices or quantum numbers, n and l. Physically acceptable solutions require that n must be greater than or equal to $l+1$. The smallest value for l is zero, so the smallest value for n is 1. The angular momentum quantum number affects the solution to the radial equation because it appears in the radial differential equation, (Equation 6.4.1).

The $R(r)$ functions that solve the radial differential Equation 6.4.1, are products of the associated Laguerre polynomials and the exponential factor, multiplied by a normalization factor ($N_{n,l}$) and $\left(\frac{r}{a_0}\right)^l$.

$$R(r) = N_{n,l}\left(\frac{r}{a_0}\right)^l L_{n,l}(r) e^{-\frac{r}{na_0}} \tag{6.4.4}$$

The decreasing exponential term overpowers the increasing polynomial term so that the overall wavefunction exhibits the desired approach to zero at large values of r. The first six radial functions are provided in Table 6.4.1 . Note that the functions in the table exhibit a dependence on Z, the atomic number of the nucleus. As discussed later in this chapter, other one electron systems have electronic states analogous to those for the hydrogen atom, and inclusion of the charge on the nucleus allows the same wavefunctions to be used for all one-electron systems. For hydrogen, $Z = 1$.

Table 6.4.3 : Radial functions for one-electron atoms and ions. Z is the atomic number of the nucleus, and $\rho = \frac{Zr}{a_0}$, where a_0 is the Bohr radius and r is the radial variable.

	n	l	$R_{n,l}(\rho)$
R_{10}	1	0	$2\left(\frac{Z}{a_0}\right)^{3/2} e^{-\rho}$
R_{20}	2	0	$\frac{1}{2\sqrt{2}}\left(\frac{Z}{a_0}\right)^{3/2}(2-\rho)e^{-\rho/2}$
R_{21}	2	1	$\frac{1}{2\sqrt{6}}\left(\frac{Z}{a_0}\right)^{3/2}\rho e^{-\rho/2}$
R_{30}	3	0	$\frac{2}{81\sqrt{3}}\left(\frac{Z}{a_0}\right)^{3/2}(27-18\rho+2\rho^2)e^{-\rho/3}$
R_{31}	3	1	$\frac{1}{81\sqrt{6}}\left(\frac{Z}{a_0}\right)^{3/2}(6\rho+\rho^2)e^{-\rho/3}$
R_{32}	3	2	$\frac{1}{81\sqrt{30}}\left(\frac{Z}{a_0}\right)^{3/2}\rho^2 e^{-\rho/3}$

The constraint that n be greater than or equal to $l+1$ also turns out to quantize the energy, producing the same quantized expression for hydrogen atom energy levels that was obtained from the Bohr model of the hydrogen atom.

$$E_n = -\frac{\mu e^4}{8\epsilon_0^2 h^2 n^2}$$

Bohr Theory vs. Schrödinger Theory

It is interesting to compare the results obtained by solving the Schrödinger equation with Bohr's model of the hydrogen atom. There are several ways in which the Schrödinger and Bohr models differ.

1. First, and perhaps most strikingly, the Schrödinger model does not produce well-defined orbits for the electron. The wavefunctions only give us the probability for the electron to be at various directions and distances from the proton.
2. Second, the quantization of angular momentum is different from that proposed by Bohr. Bohr proposed that the angular momentum is quantized in integer units of $\hbar$, while the Schrödinger model leads to an angular momentum of $\sqrt{(l(l+1)}\hbar$.
3. Third, the quantum numbers appear naturally during solution of the Schrödinger equation while Bohr had to postulate the existence of quantized energy states. Although more complex, the Schrödinger model leads to a better correspondence between theory and experiment over a range of applications that was not possible for the Bohr model.

Exercise 6.4.3

Explain how the Schrödinger equation leads to the conclusion that the angular momentum of the hydrogen atom can be zero, and explain how the existence of such states with zero angular momentum contradicts Bohr's idea that the electron is orbiting around the proton in the hydrogen atom.

Contributors and Attributions

- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")

6.4: Hydrogen Atomic Orbitals Depend upon Three Quantum Numbers is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.4: Hydrogen Atomic Orbitals Depend upon Three Quantum Numbers** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

6.5: s Orbitals are Spherically Symmetric

The hydrogen atom wavefunctions, $\psi(r,\theta,\varphi)$, are called atomic orbitals. An atomic orbital is a function that describes one electron in an atom. The wavefunction with $n = 1$, l = 0 is called the 1s orbital, and an electron that is described by this function is said to be "in" the ls orbital, i.e. have a 1s orbital state. The constraints on n, l, and m_l that are imposed during the solution of the hydrogen atom Schrödinger equation explain why there is a single 1s orbital, why there are three 2p orbitals, five 3d orbitals, etc. We will see when we consider multi-electron atoms, these constraints explain the features of the Periodic Table. In other words, the Periodic Table is a manifestation of the Schrödinger model and the physical constraints imposed to obtain the solutions to the Schrödinger equation for the hydrogen atom.

Visualizing the variation of an electronic wavefunction with r, θ, and φ is important because the absolute square of the wavefunction depicts the charge distribution (electron probability density) in an atom or molecule. The charge distribution is central to chemistry because it is related to chemical reactivity. For example, an electron deficient part of one molecule is attracted to an electron rich region of another molecule, and such interactions play a major role in chemical interactions ranging from substitution and addition reactions to protein folding and the interaction of substrates with enzymes.

We can obtain an energy and one or more wavefunctions for every value of n, the principal quantum number, by solving Schrödinger's equation for the hydrogen atom. A knowledge of the wavefunctions, or probability amplitudes ψ_n, allows us to calculate the probability distributions for the electron in any given quantum level. When n = 1, the wavefunction and the derived probability function are independent of direction and depend only on the distance r between the electron and the nucleus. In Figure 6.5.1 , we plot both ψ_1 and P_1 versus r, showing the variation in these functions as the electron is moved further and further from the nucleus in any one direction. (These and all succeeding graphs are plotted in terms of the atomic unit of length, $a_0 = 0.529 \times 10^{-8}\,cm.$)

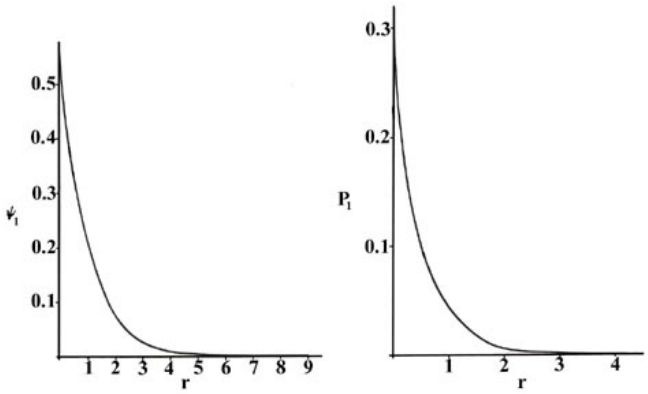


Figure 6.5.1 : The wavefunction and probability distribution as functions of r for the $n = 1$ level of the H atom. The functions and the radius r are in atomic units in this and succeeding figures.

Two interpretations can again be given to the P_1 curve. An experiment designed to detect the position of the electron with an uncertainty much less than the diameter of the atom itself (using light of short wavelength) will, if repeated a large number of times, result in Figure 6.5.1 for P_1. That is, the electron will be detected close to the nucleus most frequently and the probability of observing it at some distance from the nucleus will decrease rapidly with increasing r. The atom will be ionized in making each of these observations because the energy of the photons with a wavelength much less than 10^{-8} cm will be greater than K, the amount of energy required to ionize the hydrogen atom. If light with a wavelength comparable to the diameter of the atom is employed in the experiment, then the electron will not be excited but our knowledge of its position will be correspondingly less precise. In these experiments, in which the electron's energy is not changed, the electron will appear to be "smeared out" and we may interpret P_1 as giving the fraction of the total electronic charge to be found in every small volume element of space. (Recall that the addition of the value of P_n for every small volume element over all space adds up to unity, i.e., one electron and one electronic charge.)

Visualizing wavefunctions and charge distributions is challenging because it requires examining the behavior of a function of three variables in three-dimensional space. This visualization is made easier by considering the radial and angular parts separately, but plotting the radial and angular parts separately does not reveal the shape of an orbital very well. The shape can be revealed better in

a probability density plot. To make such a three-dimensional plot, divide space up into small volume elements, calculate $\psi^*\psi$ at the center of each volume element, and then shade, stipple or color that volume element in proportion to the magnitude of $\psi^*\psi$.

We could also represent the distribution of negative charge in the hydrogen atom in the manner used previously for the electron confined to move on a plane (Figure 6.5.2), by displaying the charge density in a plane by means of a contour map. Imagine a plane through the atom including the nucleus. The density is calculated at every point in this plane. All points having the same value for the electron density in this plane are joined by a contour line (Figure 6.5.2). Since the electron density depends only on r, the distance from the nucleus, and not on the direction in space, the contours will be circular. A contour map is useful as it indicates the "shape" of the density distribution.

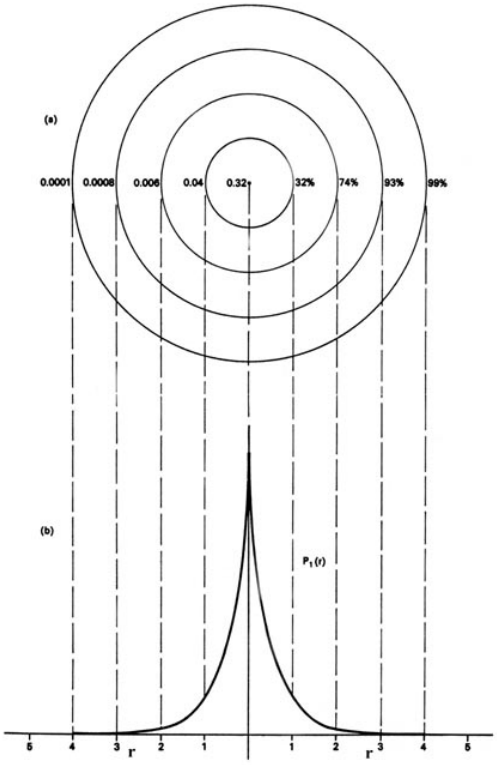


Figure 6.5.2 : (a) A contour map of the electron density distribution in a plane containing the nucleus for the $n = 1$ level of the H atom. The distance between adjacent contours is 1 au. The numbers on the left-hand side on each contour give the electron density in au. The numbers on the right-hand side give the fraction of the total electronic charge which lies within a sphere of that radius. Thus 99% of the single electronic charge of the H atom lies within a sphere of radius 4 au (or diameter = 4.2×10^{-8} cm). (b) This is a profile of the contour map along a line through the nucleus. It is, of course, the same as that given previously in Figure 6.5.1 for P_1, but now plotted from the nucleus in both directions.

When the electron is in a definite energy level we shall refer to the P_n distributions as ***electron density distributions***, since they describe the manner in which the total electronic charge is distributed in space. The electron density is expressed in terms of the number of electronic charges per unit volume of space, e^-/V. The volume V is usually expressed in atomic units of length cubed, and one atomic unit of electron density is then e^-/a_0^3. To give an idea of the order of magnitude of an atomic density unit, 1 au of charge density $e^-/a_0^3 = 6.7$ electronic charges per cubic Ångstrom. That is, a cube with a length of 0.52917×10^{-8} cm, if uniformly filled with an electronic charge density of 1 au, would contain 6.7 electronic charges.

For every value of the energy E_n, for the hydrogen atom, there is a degeneracy equal to n^2. Therefore, for $n = 1$, there is but one atomic orbital and one electron density distribution. However, for $n = 2$, there are four different atomic orbitals and four different electron density distributions, all of which possess the same value for the energy, E_2. Thus for all values of the principal quantum number n there are n^2 different ways in which the electronic charge may be distributed in three-dimensional space and still possess the same value for the energy. For every value of the principal quantum number, ***one*** of the possible atomic orbitals is independent of direction and gives a spherical electron density distribution which can be represented by circular contours as has been exemplified above for the case of $n = 1$. The other atomic orbitals for a given value of n exhibit a directional dependence and

predict density distributions which are not spherical but are concentrated in planes or along certain axes. The angular dependence of the atomic orbitals for the hydrogen atom and the shapes of the contours of the corresponding electron density distributions are intimately connected with the angular momentum possessed by the electron.

Methods for separately examining the radial portions of atomic orbitals provide useful information about the distribution of charge density within the orbitals. Graphs of the radial functions, $R(r)$, for the 1s and 2s orbitals plotted in Figure 6.5.3 . The 1s function in Figure **6.5.3**; starts with a high positive value at the nucleus and exponentially decays to essentially zero after 5 Bohr radii. The high value at the nucleus may be surprising, but as we shall see later, the probability of finding an electron at the nucleus is vanishingly small.

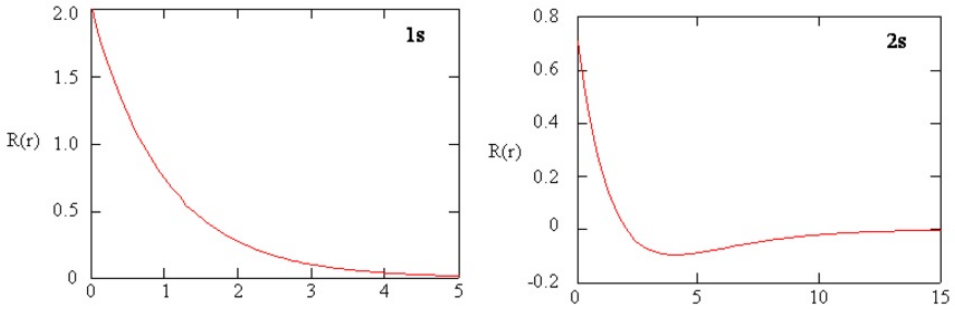


Figure 6.5.3 : Radial function, $R(r)$, for the 1s and 2s orbitals. For an interactive graph click here.

Next notice how the radial function for the 2s orbital, Figure **6.5.3**;, goes to zero and becomes negative. This behavior reveals the presence of a radial node in the function. A radial node occurs when the radial function equals zero other than at $r = 0$ or $r = \infty$. Nodes and limiting behaviors of atomic orbital functions are both useful in identifying which orbital is being described by which wavefunction. For example, all of the s functions have non-zero wavefunction values at $r = 0$.

? Exercise 6.5.1

Examine the mathematical forms of the radial wavefunctions. What feature in the functions causes some of them to go to zero at the origin while the s functions do not go to zero at the origin?

? Exercise 6.5.2

What mathematical feature of each of the radial functions controls the number of radial nodes?

? Exercise 6.5.3 : Radial Nodes

At what value of r does the 2s radial node occur?

? Exercise 6.5.4

Make a table that provides the energy, number of radial nodes, and the number of angular nodes and total number of nodes for each function with $n = 1$, $n = 2$, and $n = 3$. Identify the relationship between the energy and the number of nodes. Identify the relationship between the number of radial nodes and the number of angular nodes.

Radial probability densities for the 1s and 2s atomic orbitals are plotted in Figure 6.5.4 .

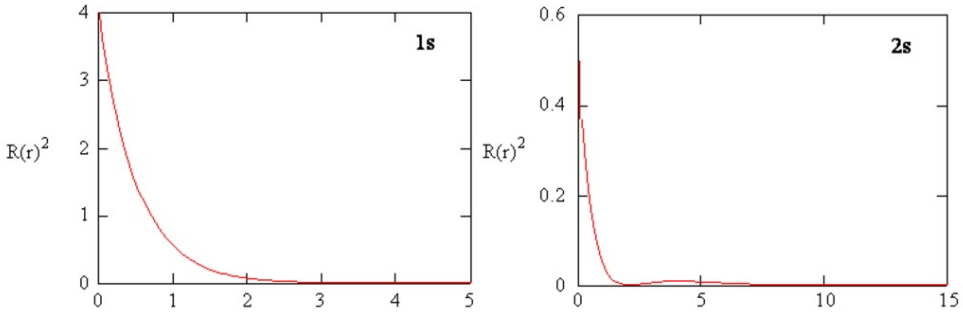


Figure 6.5.4 : Radial densities ($R(r)^*R(r)$) for the 1s and 2s orbitals.

Radial Distribution Functions

Rather than considering the amount of electronic charge in one particular small element of space, we may determine the total amount of charge lying within a thin spherical shell of space. Since the distribution is independent of direction, consider adding up all the charge density which lies within a volume of space bounded by an inner sphere of radius r and an outer concentric sphere with a radius only infinitesimally greater, say $r + \Delta r$. The area of the inner sphere is $4\pi r^2$ and the thickness of the shell is Δr. Thus the volume of the shell is $4\pi r^2 \Delta r$ and the product of this volume and the charge density $P_1(r)$, which is the charge or number of electrons per unit volume, is therefore the total amount of electronic charge lying between the spheres of radius r and $r + \Delta r$. The product $4\pi r^2 P_n$ is given a special name, the **radial distribution function**.

Volume Element for a Shell in Spherical Coordinates

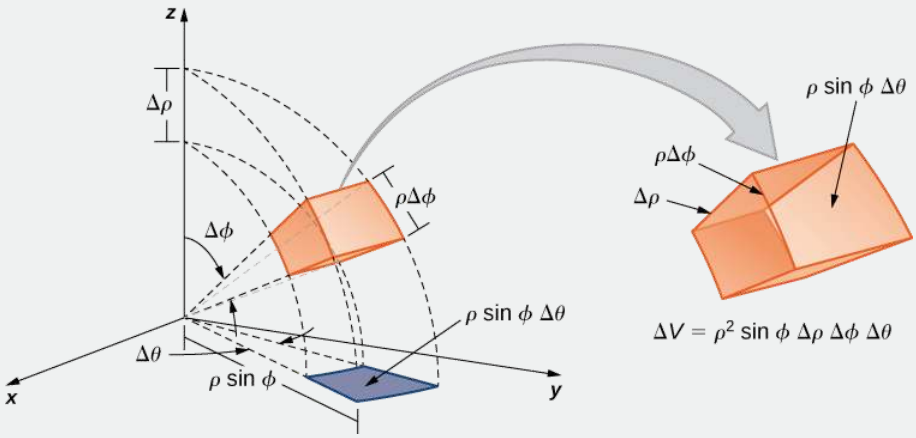
The reader may wonder why the volume of the shell is not taken as:

$$\frac{4}{3}\pi\left[(r+\Delta r)^3 - r^3\right]$$

the difference in volume between two concentric spheres. When this expression for the volume is expanded, we obtain

$$\frac{4}{3}\pi\left(3r^2\Delta r + 3r\Delta r^2 + \Delta r^3\right)$$

and for very small values of Δr the $3r\Delta r^2$ and Δr^3 terms are negligible in comparison with $3r^2\Delta r$. Thus for small values of Δr, the two expressions for the volume of the shell approach one another in value and when Δr represents an infinitesimal small increment in r they are identical.



The volume element of a box in spherical coordinates. (CC BY; OpenStax).

The radial distribution function is plotted in Figure 6.5.5 for the ground state of the hydrogen atom.

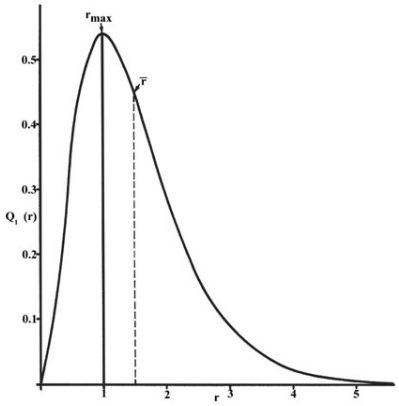


Figure 6.5.5 : The radial distribution function for an H atom. The value of this function at some value of r when multiplied by δr gives the number of electronic charges within the thin shell of space lying between spheres of radius r and $r + \delta r$.

The curve passes through zero at $r = 0$ since the surface area of a sphere of zero radius is zero. As the radius of the sphere is increased, the volume of space defined by $4\pi r^2 Dr$ increases. However, as shown in Figure 6.5.4 , the absolute value of the electron density at a given point decreases with r and the resulting curve must pass through a maximum. This maximum occurs at $r_{max} = a_0$. Thus more of the electronic charge is present at a distance a_o, out from the nucleus than at any other value of r. Since the curve is unsymmetrical, the average value of r, denoted by $\bar{r}$, is not equal to r_{max}. The average value of r is indicated on the figure by a dashed line. A "picture" of the electron density distribution for the electron in the $n = 1$ level of the hydrogen atom would be a spherical ball of charge, dense around the nucleus and becoming increasingly diffuse as the value of r is increased.

> *The radial distribution function gives the probability density for an electron to be found anywhere on the surface of a sphere located a distance r from the proton. Since the area of a spherical surface is $4\pi r^2$, the radial distribution function is given by $4\pi r^2 R(r)^* R(r)$.*

Radial distribution functions are shown in Figure 6.5.6 . At small values of r, the radial distribution function is low because the small surface area for small radii modulates the high value of the radial probability density function near the nucleus. As we increase r, the surface area associated with a given value of r increases, and the r^2 term causes the radial distribution function to increase even though the radial probability density is beginning to decrease. At large values of r, the exponential decay of the radial function outweighs the increase caused by the r^2 term and the radial distribution function decreases.

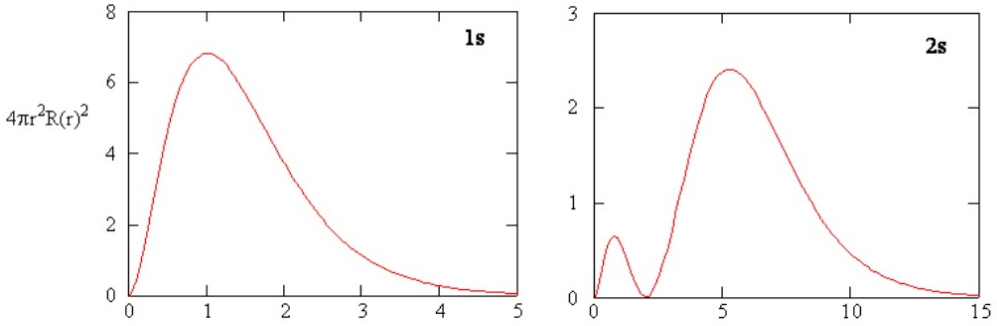


Figure 6.5.6 : The radial distribution function ($4\pi r^2 R(r)^* R(r)$) for the 1s and 2s orbitals. Compare to the radial functions in Figure 6.5.3 or the radial densities in Figure 6.5.4 . For an interactive graph click here.

Example 6.5.1 :

Calculate the probability of finding a 1s hydrogen electron being found within distance $2a_o$ from the nucleus.

Summary

This completes the description of the most stable state of the hydrogen atom, the state for which $n = 1$. Before proceeding with a discussion of the excited states of the hydrogen atom we must introduce a new term. When the energy of the electron is increased

to another of the allowed values, corresponding to a new value for n, y_n and P_n change as well. The wavefunctions y_n for the hydrogen atom are given a special name, ***atomic orbitals***, because they play such an important role in all of our future discussions of the electronic structure of atoms. In general the word orbital is the name given to a wavefunction which determines the motion of a single electron. If the one-electron wavefunction is for an atomic system, it is called an atomic orbital.

Do not confuse the word orbital with the classical word and notion of an orbit. First, an orbit implies the knowledge of a definite trajectory or path for a particle through space which in itself is not possible for an electron. Secondly, an orbital, like the wavefunction, has no physical reality but is a mathematical function which when squared gives the physically measurable electron density distribution.

Contributors and Attributions

- Dr. Richard F.W. Bader (Professor of Chemistry / McMaster University)
- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")

6.5: s Orbitals are Spherically Symmetric is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.5: s-orbitals are Spherically Symmetric** is licensed CC BY-NC-SA 4.0.

6.6: Orbital Angular Momentum and the p-orbitals

Learning Objectives

- To relate the classical orbital angular momentum for an particle to the quantum equivalent
- Characterize the mangnitude and orientation of orbital angular momentum for an electron in terms of quantum numbers

Classical Orbital Angular Momentum

The physical quantity known as angular momentum plays a dominant role in the understanding of the electronic structure of atoms. To gain a physical picture and feeling for the angular momentum it is necessary to consider a model system from the classical point of view. The simplest classical model of the hydrogen atom is one in which the electron moves in a circular orbit with a constant speed or angular velocity (Figure 6.6.1). Just as the linear momentum $m\vec{v}$ plays a dominant role in the analysis of linear motion, so angular momentum (L) plays the central role in the analysis of a system with circular motion as found in the model of the hydrogen atom.

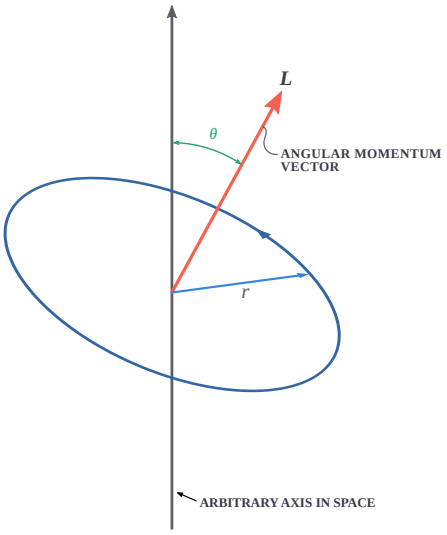


Figure 6.6.1 : The angular momentum vector for a classical model of the atom. (CC BY-NC; Ümit Kaya via LibreTexts)

In Figure 6.6.1 , m is the mass of the electron, $\vec{v}$ is the linear velocity (the velocity the electron would possess if it continued moving at a tangent to the orbit) and r is the radius of the orbit. The linear velocity $\vec{v}$ is a vector since it possesses at any instant both a magnitude and a direction in space. Obviously, as the electron rotates in the orbit the direction of $\vec{v}$ is constantly changing, and thus the linear momentum $m\vec{v}$ is ***not*** constant for the circular motion. This is so even though the speed of the electron (i.e, the magnitude of $\vec{v}$ which is denoted by $|\vec{v}|$) remains unchanged. According to Newton's second law, a force must be acting on the electron if its momentum changes with time. This is the force which prevents the electron from flying on tangent to its orbit. In an atom the attractive force which contains the electron is the electrostatic force of attraction between the nucleus and the electron, directed along the radius *r* at right angles to the direction of the electron's motion.

The angular momentum, like the linear momentum, is a vector and is defined as follows:

$$|\vec{L}| = mvr$$

The angular momentum vector $\vec{L}$ is directed along the axis of rotation. From the definition it is evident that the angular momentum vector will remain constant as long as the speed of the electron in the orbit is constant ($|\vec{v}|$ remains unchanged) and the plane and radius of the orbit remain unchanged. Thus for a given orbit, the angular momentum is constant as long as the angular velocity of the particle in the orbit is constant. In an atom the only force on the electron in the orbit is directed along r; it has no component in the direction of the motion. The force acts in such a way as to change only the linear momentum. Therefore, while the linear momentum is not constant during the circular motion, the angular momentum is. A force exerted on the particle in the direction of

the vector $\vec{v}$ would change the angular velocity and the angular momentum. When a force is applied which does change $\vec{L}$, a ***torque*** is said to be acting on the system. Thus angular momentum and torque are related in the same way as are linear momentum and force.

Quantum Angular Momentum

The important point of the above discussion is that both the angular momentum and the energy of an atom remain constant if the atom is left undisturbed. Any physical quantity which is constant in a classical system is both conserved and quantized in a quantum mechanical system. Thus both the energy and the angular momentum are quantized for an atom.

> *Any physical quantity which is constant in a classical system is both conserved and quantized in a quantum mechanical system.*

There is a quantum number, denoted by l, which governs the magnitude of the angular momentum, just as the quantum number n determines the energy. The ***magnitude*** of the angular momentum may assume only those values given by:

$$|L| = \sqrt{l(l+1)}\hbar \tag{6.6.1}$$

with $l = 0, 1, 2, 3, ...n-1$.

Furthermore, the value of *n* limits the maximum value of the angular momentum as the value of *l* cannot be greater than *n* - 1. For the state *n* = 1 discussed above, l may have the value of zero only. When *n* = 2, *l* may equal 0 or 1, and for *n* = 3, *l* = 0 or 1 or 2, etc. When *l* = 0, it is evident from Equation 6.6.1 that the angular momentum of the electron is zero. The atomic orbitals which describe these states of zero angular momentum are called *s* orbitals. The *s* orbitals are distinguished from one another by stating the value of *n*, the principal quantum number. They are referred to as the 1*s*, 2*s*, 3*s*, etc., atomic orbitals.

The preceding discussion referred to the 1s orbital since for the ground state of the hydrogen atom $n = 1$ and $l = 0$. This orbital, and all s orbitals in general, predict spherical density distributions for the electron as discussed previously.

It is common usage to refer to an electron as being "in" an orbital even though an orbital is, but a mathematical function with no physical reality. To say an electron is in a particular orbital is meant to imply that the electron is in the quantum state which is described by that orbital. For example, when the electron is in the 2*s* orbital the hydrogen atom is in a state for which $n = 2$ and $l = 0$.

Comparing these results with those for the 1*s* orbital in Figure 6.6.2 we see that as n increases the average value of r increases. This agrees with the fact that the energy of the electron also increases as n increases. The increased energy results in the electron being on the average pulled further away from the attractive force of the nucleus. As in the simple example of an electron moving on a line, nodes (values of r for which the electron density is zero) appear in the probability distributions. The number of nodes increases with increasing energy and equals $n-1$.

When the electron possesses angular momentum the density distributions are no longer spherical. In fact for each value of l, the electron density distribution assumes a characteristic shapes in Figure 6.6.2 .

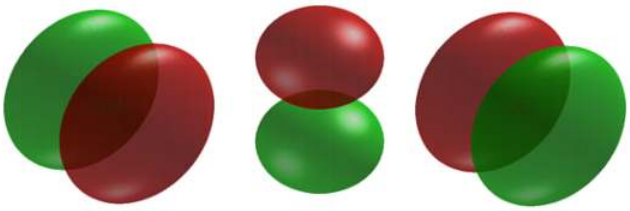


Figure 6.6.2 : The appearance of the three 2p orbitals in three-dimensional space. (CC BY-SA 3.0; I, Sarxos).

When $l = 1$, the orbitals are called *p* orbitals. In this case the orbital and its electron density are concentrated along a line (axis) in space. The 2*p* orbital or wavefunction is positive in value on one side and negative in value on the other side of a plane which is perpendicular to the axis of the orbital and passes through the nucleus. The orbital has a node in this plane, and consequently an electron in a 2*p* orbital does not place any electronic charge density at the nucleus. The electron density of a 1*s* orbital, on the other hand, is a maximum at the nucleus. The same diagram for the 2*p* density distribution is obtained for any plane which contains this

axis. Thus in three dimensions the electron density would appear to be concentrated in two lobes, one on each side of the nucleus, each lobe being circular in cross section Figure 6.6.3 .

> *An electron possesses orbital angular momentum has a density distributions is no longer spherical.*

The m_l Quantum Number and Magnetic Fields

The magnetic quantum number, designated by the letter m_l, is the third quantum numbers which describe the unique quantum state of an electron. The magnetic quantum number distinguishes the orbitals available within a subshell, and is used to calculate the azimuthal component of the orientation of the orbital in space. As with our discussion of rigid rotors, the quantum number m_l refers to the projection of the angular momentum in this arbitrarily chosen direction, conventionally called the z direction or quantization axis. L_z, the magnitude of the angular momentum in the z direction, is given by the formula

$$L_z = m_l \hbar$$

The quantum number m_l refers, loosely, to the direction of the angular momentum vector. The magnetic quantum number m_l only affects the electron's energy if it is in a magnetic field because in the absence of one, all spherical harmonics corresponding to the different arbitrary values of m_l are equivalent. The magnetic quantum number determines the energy shift of an atomic orbital due to an external magnetic field (this is called the Zeeman effect) - hence the name magnetic quantum number. However, the actual magnetic dipole moment of an electron in an atomic orbital arrives not only from the electron angular momentum, but also from the electron spin, expressed in the spin quantum number, which is the fourth quantum number. m_s and discussed in the next chapter.

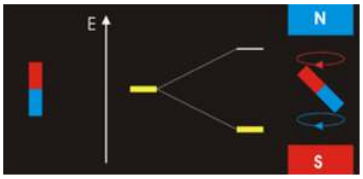


Figure 6.6.3 : The orbiting electron with a non-zero l value acts like a magnetic field with is no energetic difference for any particular orientation (only one energy state, on the left). However, in external magnetic field there is a high-energy state and a low-energy state depending on the relative orientations of the magnet to the external field. (CC SA-BY 3.0; Darekk2).

m_l Which Number Corresponds to which p-Orbital?

The answer is complicated; while $m_l = 0$ corresponds to the p_z, the orbitals for $m_l = +1$ and $m_l = -1$ lie in the xy-plane (see Spherical Harmonics), but not on the axes. The reason for this outcome is that the wavefunctions are usually formulated in spherical coordinates to make the math easier, but graphs in the Cartesian coordinates make more intuitive sense for humans. The p_x and p_y orbitals are constructed via a linear combination approach from radial and angular wavefunctions and converted into xy (this was discussed previously). Thus, it is not possible to directly correlate the values of $m_l = \pm 1$ with specific orbitals. The notion that we can do so is sometimes presented in introductory courses to make a complex mathematical model just a little bit simpler and more intuitive, but it is incorrect.

The three wavefunctions for $n = 2$ and $l = 1$ are as follows.

$$|\psi_{2,1,0}\rangle = r\cos\theta R(r) \tag{6.6.2}$$

$$|\psi_{2,1,+1}\rangle = -\frac{r}{2}\sin\theta e^{i\phi} R(r) \tag{6.6.3}$$

$$|\psi_{2,1,-1}\rangle = +\frac{r}{2}\sin\theta e^{-i\phi} R(r) \tag{6.6.4}$$

The notation is $|\psi_{n,l,m_l}\rangle$ with $R(r)$ is the radial component of this wavefuction, θ is the angle with respect to the z-axis and ϕ is the angle with respect to the xz-plane.

$$R(r) = \sqrt{\frac{Z^5}{32\pi a_0^5}} e^{-Zr/2a_0}$$

in which Z is the atomic number (or probably better nuclear charge) and a_0 is the Bohr radius.

In switching from spherical to Cartesian coordinates, we make the substitution $z = r\cos\theta$, so:

$$|\psi_{2,1,0}\rangle = zR(r)$$

This is ψ_{2p_z} since the value of ψ is dependent on z: when $z = 0$; $\psi = 0$, which is expected since $z = 0$ describes the xy-plane.

The other two wavefunctions are degenerate in the xy-plane. An equivalent statement is that these two orbitals do not lie on the x- and y-axes, but rather bisect them. Thus it is typical to take linear combinations of them to make the equation look prettier. If any set of wavefunctions is a solution to the Schrödinger equation, then any set of linear combinations of these wavefunctions must also be a solution (Section 2.4). We can do this because of the linearity of the Schrödinger equation.

In the equations below, we're going to make use of some trigonometry, notably Euler's formula:

$$e^{i\phi} = \cos\phi + i\sin\phi \tag{6.6.5}$$

$$\sin\phi = \frac{e^{i\phi} - e^{-i\phi}}{2i} \tag{6.6.6}$$

$$\cos\phi = \frac{e^{i\phi} + e^{-i\phi}}{2} \tag{6.6.7}$$

We're also going to use $x = \sin\theta\cos\phi$ and $y = \sin\theta\sin\phi$.

$$\begin{aligned}
\psi_{2p_x} &= \frac{1}{\sqrt{2}}\left(\psi_{2,1,+1} - \psi_{2,1,-1}\right) \\
&= \frac{1}{2}\left(e^{i\phi} + e^{-i\phi}\right) r\sin\theta f(r) \\
&= r\sin\theta\cos\phi f(r) = x f(r) \\
\psi_{2p_y} &= \frac{i}{\sqrt{2}}\left(\psi_{2,1,+1} + \psi_{2,1,-1}\right) \\
&= \frac{1}{2i}\left(e^{i\phi} - e^{-i\phi}\right) r\sin\theta f(r) \\
&= r\sin\theta\sin\phi f(r) = y f(r)
\end{aligned}$$

So, while $m_l = 0$ corresponds to $|p_z\rangle$, $m_l = +1$ and $m_l = -1$ **cannot** be directly assigned to either $|p_x\rangle$ or $|p_y\rangle$, but rather a combination of $|p_x\rangle$ and $|p_y\rangle$. An alternative description is that $m_l = +1$ might correspond to $(|p_x\rangle + |p_y\rangle)$ and $m_l = -1$ might correspond to $(|p_x\rangle - |p_y\rangle)$.

d-Orbitals (even higher angular momenta wavefunctions)

When $l = 2$, the orbitals are called *d* orbitals and Figure 6.6.4 shows the contours in a plane for a 3*d* orbital and its density distribution. Notice that the density is again zero at the nucleus and that there are now two nodes in the orbital and in its density distribution. As the angular momentum of the electron increases, the density distribution becomes increasingly concentrated along an axis or in a plane in space. Only electrons in s orbitals with **zero angular momentum** give spherical density distributions and in addition place charge density at the position of the nucleus.



Figure 6.6.4 : The appearance of the 3d electron density distribution in three-dimensional space. (CC BY-SA 3.0; I, Sarxos)

As with the p-orbitals, the only d-orbital that a specific m_l can be ascribed is the d_{z^2} orbitals with $m_l = 0$. The rest are linear combinations of the hydrogen atom wavefunctions with complex spherical harmonic angular components.

There seems to be neither rhyme nor reason for the naming of the states corresponding to the different values of ℓ (s, *p*, *d*, *f* for *l* = 0, 1, 2, 3). This set of labels had its origin in the early work of experimental atomic spectroscopy. The letter *s* stood for

sharp, *p* for principal, *d* for diffuse and *f* for fundamental in characterizing spectral lines. From the letter *f* onwards the naming of the orbitals is alphabetical $l = 4, 5, 6 \rightarrow g, h, i, ...$

We have not as yet accounted for the full degeneracy of the hydrogen atom orbitals which we stated earlier to be n^2 for every value of n. For example, when $n = 2$, there are four distinct atomic orbitals. The remaining degeneracy is again determined by the angular momentum of the system. Since angular momentum like linear momentum is a vector quantity, we may refer to the component of the angular momentum vector which lies along some chosen axis. For reasons we shall investigate, the number of values a particular component can assume for a given value of l is $(2l + 1)$. Thus when $l = 0$, there is no angular momentum and there is but a single orbital, an s orbital. When $l = 1$, there are three possible values for the component $(2 \times 1 + 1)$ of the total angular momentum which are physically distinguishable from one another. There are, therefore, three p orbitals. Similarly there are five d orbitals, $(2 \times 2 + 1)$, seven f orbitals, $(2 \times 3 + 1)$, etc. All of the orbitals with the same value of n and l, the three 2p orbitals for example, are similar but differ in their spatial orientations.

To gain a better understanding of this final element of degeneracy, we must consider in more detail what quantum mechanics predicts concerning the angular momentum of an electron in an atom.

Contributors and Attributions

- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")
- Dr. Richard F.W. Bader (Professor of Chemistry / McMaster University)
- Stack Exchange (Ben Norris and Loong)

6.6: Orbital Angular Momentum and the p-orbitals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.6: Orbital Angular Momentum and the p-Orbitals** is licensed CC BY-NC-SA 4.0.

6.7: The Helium Atom Cannot Be Solved Exactly

Learning Objectives

- Adding electrons to the quantum hydrogen atom results in analytically unsolvable Schrödinger Equations (they exist, we just do not have analytical forms for them)
- A basic aspect of the corresponding multi-electron Hamiltonians is that they are NOT separable with respect to the spatial coordinate of each electron
- The solutions to multi-electron Schrödinger Equations are called multi-electron wavefunctions and they are often approximated as a product of single-electron wavefunctions (called the orbital approximation).

Multi-electron Hamiltonians

The second element in the periodic table provides our first example of a quantum-mechanical problem which *cannot* be solved exactly. Nevertheless, as we will show, approximation methods applied to helium can give accurate solutions in perfect agreement with experimental results. In this sense, it can be concluded that quantum mechanics is correct for atoms more complicated than hydrogen. By contrast, the Bohr theory failed miserably in attempts to apply it beyond the hydrogen atom.

Figure 6.7.1 shows a schematic representation of a helium atom with two electrons whose coordinates are given by the vectors $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$. The electrons are separated by a distance $\boldsymbol{r_{12} = |r_1 - r_2|}$. The origin of the coordinate system is fixed at the nucleus. As with the hydrogen atom, the nuclei for multi-electron atoms are so much heavier than an electron that the nucleus is assumed to be the center of mass. Fixing the origin of the coordinate system at the nucleus allows us to exclude translational motion of the center of mass from our quantum mechanical treatment.

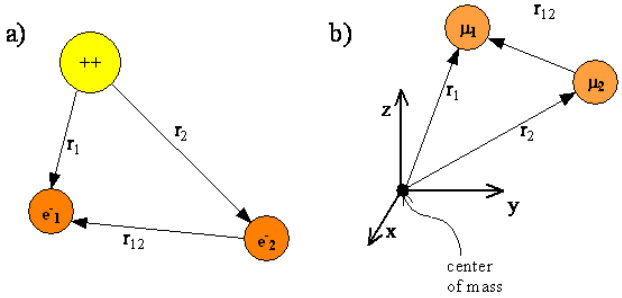


Figure 6.7.1 : (a) The nucleus (++) and electrons (e-) of the helium atom. (b) Equivalent reduced particles with the center of mass (approximately located at the nucleus) at the origin of the coordinate system. Note that $\boldsymbol{\mu_1}$ and $\boldsymbol{\mu_2 \approx m_e}$.

The Hamiltonian operator for the hydrogen atom serves as a reference point for writing the Hamiltonian operator for atoms with more than one electron. Start with the same general form we used for the hydrogen atom Hamiltonian

$$\hat{H} = \hat{T} + \hat{V} \tag{6.7.1}$$

Include a kinetic energy term for each electron and a potential energy term for the attraction of each negatively charged electron for the positively charged nucleus and a potential energy term for the mutual repulsion of each pair of negatively charged electrons. The He atom Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m_e}(\nabla_1^2 + \nabla_2^2) + V_1(r_1) + V_2(r_2) + V_{12}(r_{12}) \tag{6.7.2}$$

where

$$V_1(r_1) = -\frac{2e^2}{4\pi\epsilon_0 r_1} \tag{6.7.3}$$

$$V_2(r_2) = -\frac{2e^2}{4\pi\epsilon_0 r_2} \tag{6.7.4}$$

$$V_{12}(r_{12}) = \frac{e^2}{4\pi\epsilon_0 r_{12}} \tag{6.7.5}$$

The two-electron Hamiltonian in Equation 6.7.2 can be extended to any atom or ion by replacing the He nuclear charge of +2 with a general charge Z; e.g.

$$V_1(r_1) = -\frac{Ze^2}{4\pi\epsilon_0 r_1} \tag{6.7.6}$$

and including terms for the additional electrons. The subsequent multi-electron atom with n electron is

$$\hat{H} = \underbrace{-\frac{\hbar^2}{2m_e}\sum_i^n \nabla_i^2}_{\text{Kinetic Energy}} + \underbrace{\sum_i^n V_i(r_i)}_{\text{Coulombic Attraction}} + \underbrace{\sum_{i\neq j}^{n,n} V_{ij}(r_{ij})}_{\text{electron-electron Repulsion}} \tag{6.7.7}$$

This multi-electron Hamiltonian is qualitatively similar to the 2-electron Hamiltonian (Equation 6.7.1) with each electron having its own kinetic energy and nuclear potential energy terms (Equations 6.7.3 and 6.7.4). The other big difference between single electron systems and multi-electron systems is the presence of the $V_{ij}(r_{ij})$ terms which contain $1/r_{ij}$, where r_{ij} is the distance between electrons i and j. These terms account for the electron-electron repulsion that we expect between like-charged particles.

? Exercise 6.7.1 : Multi-electron atom Hamiltonians

For the generalized multi-electron atom Hamiltonian (Equation 6.7.7):

a. Explain the origin of each of the three summations.
b. What do these summations over (i.e., what is the origin of the summing index)?
c. Write expressions for $V_i(r_i)$ and $V_{ij}(r_{ij})$.

? Exercise 6.7.2 : Boron Atom

Boron is the fifth element of the periodic table (Z=5) and is located in Group 13.

a. Write the multi-electron Hamiltonian for a $^{11}\mathbf{B}$ atom.
b. Would it be any different for a $^{11}\mathbf{B}^+$ ion?
c. Would it be any different for a $^{10}\mathbf{B}$ atom?

Multi-electron Wavefunctions and the Orbital Approximation

Given what we have learned from the previous quantum mechanical systems we've studied, we predict that exact solutions to the multi-electron Schrödinger equation in Equation 6.7.7 would consist of a family of multi-electron wavefunctions, each with an associated energy eigenvalue. These wavefunctions and energies would describe the ground and excited states of the multi-electron atom, just as the hydrogen wavefunctions and their associated energies describe the ground and excited states of the hydrogen atom. We would predict quantum numbers to be involved, as well.

The fact that electrons interact through their Coulomb repulsion means that an exact wavefunction for a multi-electron system would be a single function that depends *simultaneously* upon the coordinates of all the electrons; i.e., **a multi-electron wavefunction**,

$$\psi(\vec{r}_1, \vec{r}_2, \vec{r}_3, \vec{r}_4, \cdots \vec{r}_n)$$

The modulus squared of such a wavefunction would describe the probability of finding the electrons (though not specific ones) at a designated volume (V) in the atom.

$$p(V) = \int_V |\psi(\vec{r}_1, \vec{r}_2, \vec{r}_3, \vec{r}_4, \cdots \vec{r}_n)|^2 d\tau$$

All of the electrons are described simultaneously by a multi-electron wavefunction, so the total amount of electron density represented by the wavefunction equals the number of electrons in the atom.

$$\int_{\text{all space}} |\psi(\vec{r}_1, \vec{r}_2, \vec{r}_3, \vec{r}_4, \cdots \vec{r}_i)|^2 d\tau = n$$

Unfortunately, the Coulomb repulsion terms (Equation 6.7.5) make it impossible to find an exact solution to the Schrödinger equation for many-electron atoms and molecules even for two electrons atoms. We have to rely on approximations and the *orbital approximation* is central to basic chemistry concepts.

The Orbital Approximation

The most basic approximations to the exact solutions to a multi-electron atom Hamiltonian, $\hat{H}$, (Equation 6.7.7) involve writing a multi-electron wavefunction ($\psi(r_1, r_2, \cdots, r_n)$) as a simple product of single-electron wavefunctions ($\varphi_i(r_i)$):

$$\psi(r_1, r_2, \cdots, r_n) \approx \varphi_1(r_1)\varphi_2(r_2)\cdots\varphi_n(r_n) \tag{6.7.8}$$

or in Dirac notation

$$|\psi(r_1, r_2, \cdots, r_n)\rangle \approx |\varphi_1(r_1)\rangle|\varphi_2(r_2)\rangle\cdots|\varphi_n(r_n)\rangle \tag{6.7.9}$$

The energy of the atom in the state associated with a specific multi-electron wavefunction (E) is obtained from the multi-electron Schrödinger Equation

$$\hat{H}\psi(r_1, r_2, \cdots, r_n) = E\psi(r_1, r_2, \cdots, r_n)$$

Within the approximation in Equation 6.7.8, E can be expressed sum of the energies of the one-electron components (ϵ_i).

$$E \approx \sum_i \epsilon_i$$

This is called the **orbital approximation**.

By writing the multi-electron wavefunction as a product of single-electron functions (Equations 6.7.8 or 6.7.9), we conceptually transform a multi-electron atom into a collection of individual electrons located in individual orbitals whose spatial characteristics and energies can be separately identified. For atoms, these single-electron wavefunctions are called *atomic orbitals* and resemble the wavefunctions for hydrogen-like atoms. For molecules, as we will see in the following chapters, these are called *molecular orbitals*. While a great deal can be learned from such an analysis, it is important to keep in mind that such a discrete, compartmentalized picture of the electrons is an approximation, albeit a powerful one.

6.7: The Helium Atom Cannot Be Solved Exactly is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.7: The Helium Atom Cannot Be Solved Exactly** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0. Original source: https://web.archive.org/web/20200619182410/http://www.chemeddl.org.

6.E: The Hydrogen Atom (Exercises)

Solutions to select questions can be found online.

6.5

$$\int_{-1}^{1} T_n(x)T_m(x)\frac{1}{\sqrt{1-x^2}}dx = \begin{cases} 0, & n \neq m \\ \pi, & n = m = 0 \\ \pi/2, & n = m \neq 0 \end{cases}$$

First 6 Chebyshev Polynomials

$$T_0(x) = 1$$
$$T_1(x) = x$$
$$T_2(x) = 2x^2 - 1$$
$$T_3(x) = 4x^3 - 3x$$
$$T_4(x) = 8x^4 - 8x^2 + 1$$
$$T_5(x) = 16x^5 - 20x^3 + 5x$$

Use the orthogonality of Chebyshev polynomials to determine what the following polynomials are equal to

a. $\int_{-1}^{1} x^2 \frac{dx}{\sqrt{1-x^2}}$
b. $\int_{-1}^{1} 4x^3 - 2x \frac{dx}{\sqrt{1-x^2}}$
c. $\int_{-1}^{1} 1 \frac{dx}{\sqrt{1-x^2}}$
d. $\int_{-1}^{1} 4x^4 - 4x^2 + 1 \frac{dx}{\sqrt{1-x^2}}$

6.6

Use Eq. 6.47 to generate the radial functions $R_{nl}(r)$ for $n = 1, 2$.

6.29

Compare ψ_{310} and ψ_{311}.

Hint: What do the subscripts tell you about the wave function? What do they denote?

6.30

What is the probability density of the 3p orbital by evaluating

$$\left(\sum_{m=-1}^{1} \psi_{31m}^2\right)$$

6.34

Find the energy, and wavefunction for a single electron located in the 2p orbital of the hydrogen atom. Include all possible wavefunctions.

6.37

The Hamiltonian is given by $\hat{H} = \frac{-\hbar}{2m}\nabla^2 + V$ is an Hermitian Operator. Using this fact, show that

$$\int \psi^* [\hat{H}, \hat{A}] \psi d\tau = 0$$

where $\hat{A}$ is any operator.

6.38

Prove that $\langle \hat{K} \rangle = \langle V \rangle = E/2$ for a harmonic oscillator using the virial theorem

6.41

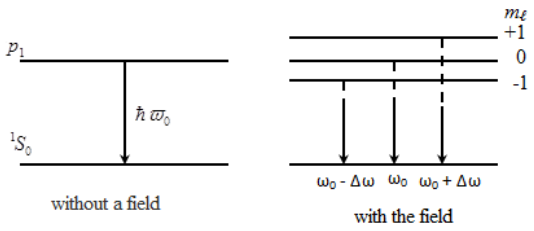
Find the expected values of $1/r$ and $1/r^2$ for a hydrogenlike atom in the $2p_z$ orbital.

6.43

Derive the classical magnetic moment of an electron orbiting a nucleus in terms of charge, mass and angular momentum.

6.46

Find the magnitude of the splitting shown in figure below. The magnetic field in the figure is at 20 T.



6.47

Consider the transition between the $l = 1$ and the $l = 2$ states for atomic hydrogen. Determine the total number of possible allowed transitions between these two states in an external magnetic field given the following selection rules

a. Light whose electric field vector is parallel to the external magnetic field's direction has a selection rule of $\Delta m = 0$ for allowed transitions.
b. Light whose electric field vector is perpendicular to the external magnetic field's direction has a selection rule of $\Delta m = \pm 1$ for allowed transitions.

6.49

Prove that

$$\hat{L_+}\hat{L_-} - \hat{L_-}\hat{L_+} = 2\hbar\hat{L_z}$$

given that

$$\hat{L_+} = \hat{L_x} + i\hat{L_y}$$

and

$$\hat{L_-} = \hat{L_x} - i\hat{L_y}.$$

6.49

Show that the commutative property applies to

$$\hat{L}_-\hat{L}_+$$

Q7.29

Calculate the ground-state energy for particle in the box model using variational method.

6.50

If two functions commute, they have mutual eigenfunctions, such as $\hat{L}^2$ and $\hat{L}_z$. These mutual eigenfunctions are also known as spherical harmonics, $Y_l^m(\theta, \phi)$, however this information is not pertinent in this case. Let

$\psi_{\alpha\beta}$ be a mutual eigenfunction of $\hat{L}^2$ and $\hat{L}_z$ so that

$$\hat{L}_z \psi_{\alpha\beta} = \beta^2 \backslash_{\alpha\beta}$$

and

$$\hat{L}_z \psi_{\alpha\beta} = \alpha\psi_{\alpha\beta}$$

Now let

$$\psi_{\alpha\beta}{}^{+1} = \hat{L}_+ \psi_{\alpha\beta}$$

Show that

$$\hat{L}_z \psi_{\alpha\beta}{}^{+1} = (\alpha + \hbar)\psi_{\alpha\beta}{}^{+1}$$

and

$$\hat{L}^2 \psi_{\alpha\beta}{}^{+1} = \beta^2 \psi_{\alpha\beta}{}^{+1}$$

This proves that if α is an eigenvalue of $\hat{L}_z$, then $\alpha + \hbar$ also is an eigenvalue.

6.E: The Hydrogen Atom (Exercises) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.E: The Hydrogen Atom (Exercises)** is licensed CC BY-NC-SA 4.0.

CHAPTER OVERVIEW

7: Approximation Methods

The Schrödinger equation for realistic systems quickly becomes unwieldy, and analytical solutions are only available for very simple systems - the ones we have described as *fundamental systems* in this module. Numerical approaches can cope with more complex problems, but are still (and will remain for a good while) limited by the available computer power. Approximations are necessary to cope with real systems. Within limits, we can use a pick and mix approach, *i.e.* use *linear combinations* of solutions of the fundamental systems to build up something akin to the real system. There are two mathematical techniques, *perturbation* and *variation* theory, which can provide a good approximation along with an estimate of its accuracy. These two approximation techniques are described in this chapter.

7.0: Slater's Rules for Shielding

7.1: The Variational Method

7.2: Linear Variational Method and the Secular Determinant

7.3: Trial Functions Can Be Linear Combinations of Functions with Variational Parameters

7.4: Perturbation Theory Expresses the Solutions in Terms of Solved Problems

7.E: Approximation Methods (Exercises)

7: Approximation Methods is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

7.0: Slater's Rules for Shielding

Learning Objective

- To quantify the shielding effect experienced by atomic electrons.

We have previously described the concepts of electron shielding, orbital penetration and effective nuclear charge, but we did so in a qualitative manner. In this section, we explore one model for quantitatively estimating the impact of electron shielding, and then use that to calculate the effective nuclear charge experienced by an electron in an atom. The model we will use is known as Slater's Rules (J.C. Slater, *Phys Rev* **1930**, *36*, 57).

Slater's Rules

The general principle behind Slater's Rule is that the actual charge felt by an electron is equal to what you'd expect the charge to be from a certain number of protons, but minus a certain amount of charge from other electrons. Slater's rules allow you to estimate the effective nuclear charge Z_{eff} from the real number of protons in the nucleus and the effective shielding of electrons in each orbital "shell" (e.g., to compare the effective nuclear charge and shielding 3d and 4s in transition metals). Slater's rules are fairly simple and produce fairly accurate predictions of things like the electron configurations and ionization energies.

Slater's Rules

- **Step 1**: Write the electron configuration of the atom in the following form:

  (1s) (2s, 2p) (3s, 3p) (3d) (4s, 4p) (4d) (4f) (5s, 5p) . . .

- **Step 2**: Identify the electron of interest, and ignore all electrons in higher groups (to the right in the list from Step 1). These do not shield electrons in lower groups
- **Step 3**: Slater's Rules is now broken into two cases:
  - the shielding experienced by an s- or p- electron,
    - electrons within same group shield **0.35**, except the 1*s* which shield **0.30**
    - electrons within the n-1 group shield **0.85**
    - electrons within the n-2 or lower groups shield **1.00**
  - the shielding experienced by n*d* or n*f* valence electrons
    - electrons within same group shield **0.35**
    - electrons within the lower groups shield **1.00**

These rules are summarized in Figure **7.0.1** and Table **7.0.1**.

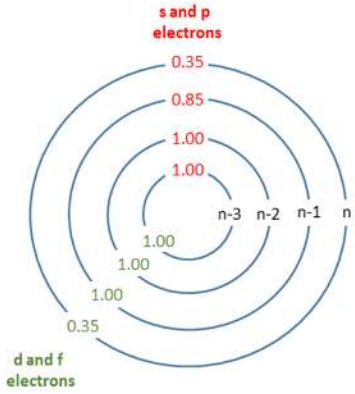


Figure **7.0.1**: Graphical depiction of Slater's rules with shielding constants indicated.

Shielding happens when electrons in lower valence shells (or the same valence shell) provide a repulsive force to valence electrons, thereby "negating" some of the attractive force from the positive nucleus. Electrons really close to the atom (n-2 or lower) pretty

much just look like protons, so they completely negate. As electrons get closer to the electron of interest, some more complex interactions happen that reduce this shielding.

Table **7.0.1**: Slater's Rules for calculating shielding constants

Group	Other electrons in the same group	Electrons in group(s) with principal quantum number n and azimuthal quantum number < l	Electrons in group(s) with principal quantum number n-1	Electrons in all group(s) with principal quantum number < n-1
[1s]	0.30	-	-	-
[*n*s,*n*p]	0.35	-	0.85	1
[*n*d] or [*n*f]	0.35	1	1	1

The shielding numbers in Table **7.0.1** were derived semi-empirically (i.e., derived from experiments) as opposed to theoretical calculations. This is because quantum mechanics makes calculating shielding effects quite difficult, which is outside the scope of this Module.

Calculating *S*

Sum together the contributions as described in the appropriate rule above to obtain an estimate of the shielding constant, S, which is found by totaling the screening by **all electrons** except the one in question.

$$S = \sum n_i S_i \tag{7.0.1}$$

where

- n_i is the number of electrons in a specific shell and subshell and
- S_i is the shielding of the electrons subject to Slater's rules (Table **7.0.1**)

Example 7.0.1: The Shielding of 3*p* Electrons of Nitrogen Atoms

What is the shielding constant experienced by a 2*p* electron in the nitrogen atom?

Given:

Nitrogen (N)

Asked for:

S, the shielding constant, for a 2*p* electron (Equation 7.0.1)

Strategy:

A. Determine the electron configuration of nitrogen, then write it in the appropriate form.

B. Use the appropriate Slater Rule to calculate the shielding constant for the electron.

Step A

N: $1s^2\ 2s^2\ 2p^3$

N: $(1s^2)(2s^2,2p^3)$

Step B

$$S[2p] = \underbrace{0.85(2)}_{\text{the 1s electrons}} + \underbrace{0.35(4)}_{\text{the 2s and 2p electrons}} = 3.10$$

As Table **7.0.1** indicates,

- the 1s electrons shield the other 2p electron to 0.85 "charges".
- the 2s and 2p electrons shield the other 2p electron equally at 0.35 "charges".

Exercise 7.0.1: The Shielding of valence *p* Electrons of Bromine Atoms

What is the shielding constant experienced by a valence *p*-electron in the bromine atom?

Example 7.0.2: The Shielding of 3*d* Electrons of Bromine Atoms

What is the shielding constant experienced by a 3*d* electron in the bromine atom?

Given:

Bromine (Br)

Asked for:

S, the shielding constant, for a 3*d* electron

Strategy:

A. Determine the electron configuration of bromine, then write it in the appropriate form.
B. Use the appropriate Slater Rule to calculate the shielding constant for the electron.

Step A

Br: $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 4s^2\ 3d^{10}\ 4p^5$

Br: $(1s^2)(2s^2,2p^6)(3s^2,3p^6)(3d^{10})(4s^2,4p^5)$

Ignore the group to the right of the 3d electrons. These do not contribute to the shielding constant.

Step B

$$S[3d] = 1.00(18) + 0.35(9) = 21.15$$

Exercise 7.0.2: The Shielding of 3*d* Electrons of Copper Atoms

What is the shielding constant experienced by a valence *d*-electron in the copper atom?

Calculating Effective Nuclear Charge

Previously, we described Z_{eff} as being less than the actual nuclear charge (Z) because of the repulsive interaction between core and valence electrons. We can quantitatively represent this difference between Z and Z_{eff} as follows:

$$S = Z - Z_{eff} \tag{7.0.2}$$

Rearranging this formula to solve for Z_{eff} we obtain:

$$Z_{eff} = Z - S \tag{7.0.3}$$

We can then substitute the shielding constant obtained using Equation 7.0.3 to calculate an estimate of Z_{eff} for the corresponding atomic electron.

Example 7.0.3: The Effective Charge of p Electrons of Boron Atoms

What is the effective nuclear charge experienced by a valence *p*- electron in boron?

Given:

Boron (B)

Asked for:

Z_{eff} for a valence *p*- electron

Strategy:

A. Determine the electron configuration of boron and identify the electron of interest.
B. Use the appropriate Slater Rule to calculate the shielding constant for the electron.

C. Use the Periodic Table to determine the actual nuclear charge for boron.
D. Determine the effective nuclear constant.

Step A

B: $1s^2\ 2s^2\ 2p^1$. The valence p- electron in boron resides in the $2p$ subshell.

B: $(1s^2)(2s^2,2p^1)$

Step B

$S[2p] = 1.00(0) + 0.85(2) + 0.35(2) = 2.40$

Step C

$Z = 5$

Step D

Using Equation 7.0.3,

$$\begin{aligned} Z_{eff} &= Z - S \\ &= 5 - 2.50 \\ &= 2.60 \end{aligned}$$

Exercise 7.0.3

What is the effective nuclear charge experienced by a valence d-electron in copper?

Summary

Slater's Rules can be used as a model of shielding. This permits us to quantify both the amount of shielding experienced by an electron and the resulting effective nuclear charge. Others performed better optimizations of Z_{eff} using variational Hartree-Fock methods. For example, Clementi and Raimondi published "Atomic Screening Constants from SCF Functions." J Chem Phys (1963) 38, 2686–2689.

References

- James L. Reed, "The Genius of Slater's Rules" , J. Chem. Educ., 1999, 76 (6), p 802
- David Tudela, "Slater's rules and electron configurations", J. Chem. Educ., 1993, 70 (11), p 956
- Kimberley A. Waldron, Erin M. Fehringer, Amy E. Streeb, Jennifer E. Trosky and Joshua J. Pearson, "Screening Percentages Based on Slater Effective Nuclear Charge as a Versatile Tool for Teaching Periodic Trends", J. Chem. Educ., 2001, 78 (5), p 635

7.0: Slater's Rules for Shielding is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.6: Slater's Rules** by Brett McCollum is licensed CC BY-NC-SA 4.0.

7.1: The Variational Method

Learning Objectives

- Appreciate the complexity of solving muliti-electron atoms
- Characterize multi-electron interactions within shielding and penetration concepts
- Use the variational method as an approximation to study insolvable problems
- User variational method to evaluate the effective nuclear charge for a specific atom

In this section we introduce the powerful and versatile variational method and use it to improve the approximate solutions we found for the helium atom using the independent electron approximation.

The True (Experimentally Determined) Energy of the Helium Atom

The helium atom has two electrons bound to a nucleus with charge $Z = 2$. The *successive* removal of the two electrons can be considered stepwise:

$$\mathrm{He} \xrightarrow{I_1} \mathrm{He}^+ + e^- \xrightarrow{I_2} \mathrm{He}^{++} + 2e^- \tag{7.1.1}$$

The *first ionization energy* I_1 is the minimum energy required to remove the first electron from helium gas and is experimentally determined:

$$I_1 = -E_{1s}(\mathrm{He}) = 24.59\ eV$$

While the second ionization energy, I_2 can be experimentally determined, it can also be calculated exactly from the hydrogen atom solutions since He^+ is a hydrogen-like ion with $Z = 2$. Hence, we have

$$\begin{aligned} I_2 &= -E_{1s}(\mathrm{He}^+) \\ &= \frac{Z^2}{2n^2} \\ &= 54.42\ \mathrm{eV} \end{aligned}$$

The energy of the three separated particles on the right side of Equation 7.1.1 is zero (by definition). Therefore the ground-state energy of helium atom is given by

$$\begin{aligned} E_{true} &= -(I_1 + I_2) \\ &= -79.02\ \mathrm{eV}. \end{aligned}$$

which can be expressed in terms of the ***Rydberg*** constant ($R_H = 13.6\ eV$) that also describes the lowest energy of the hydrogen atom

$$E_{true} = -5.8066\, R_H$$

We will attempt to reproduce this **true value**, as close as possible, by different theoretical approaches (all approximations).

The "Ignorance is Bliss" Approximation

The Hamiltonian for the Helium atom is:

$$\hat{H} = -\frac{\hbar^2}{2m_e}\nabla^2_{el_1} - \frac{\hbar^2}{2m_e}\nabla^2_{el_2} - \frac{Ze^2}{4\pi\epsilon_0 r_1} - \frac{Ze^2}{4\pi\epsilon_0 r_2} + \cancel{\frac{e^2}{4\pi\epsilon_0 r_{12}}} \tag{7.1.2}$$

If we simply *ignore* the electron-electron repulsion term, then Equation 7.1.2 can be simplified to

$$\hat{H} \approx -\frac{\hbar^2}{2m_e}\nabla^2_{el_1} - \frac{Ze^2}{4\pi\epsilon_0 r_1} - \frac{\hbar^2}{2m_e}\nabla^2_{el_2} - \frac{Ze^2}{4\pi\epsilon_0 r_2} \tag{7.1.3}$$

$$\approx h_1(r_1) + h_2(r_2) \tag{7.1.4}$$

where h_1 and h_2 are one electron Hamiltonians for electron 1 and 2, respectively, and are just the hydrogen-like Hamiltonians. The approximation in Equation 7.1.4 is convenient since electron 1 is *separable* from electron 2, so that the two-electron wavefunction is approximated as a product to two one-electron wavefunctions:

$$\Psi_{total} = \psi_{el_1}\psi_{el_2} \tag{7.1.5}$$

or in braket notation

$$|\Psi_{total}\rangle = \hat{H}|\psi_{el_1}\psi_{el_2}\rangle \tag{7.1.6}$$

With some operator algebra, something important arises - the one electron energies are **additive**:

$$\hat{H}\Psi_{total} = (\hat{H}_{el_1} + \hat{H}_{el_2})\psi_{n\ el_1}\psi_{n\ el_2} = (E_{n_1} + E_{n_2})\psi_{n\ el_1}\psi_{n\ el_2}$$

or in bra-ket notation

$$\begin{aligned}\hat{H}|\Psi_{total}\rangle &= \hat{H}|\psi_{el_1}\psi_{el_2}\rangle \\ &= (E_{n_1} + E_{n_2})|\psi_1\psi_2\rangle\end{aligned}$$

The energy for a ground state Helium atom (both electrons in lowest state) is then

$$\begin{aligned}E_{He_{1s}} &= \underbrace{E_{n_1}}_{\text{energy of single electron in helium}} + \underbrace{E_{n_2}}_{\text{energy of single electron in helium}} \\ &= -R_H\left(\frac{Z^2}{1}\right) - R_H\left(\frac{Z^2}{1}\right) \\ &= -8R_H\end{aligned}$$

This approximation significantly overestimates the true energy of the helium atom $E_{He_{1s}} = -5.8066\ R$. This is a poor approximation and we need to address electron-electron repulsion properly (or better at least).

Shielding and Penetration

One way to take electron-electron repulsion into account is to modify the form of the wavefunction. A logical modification is to change the nuclear charge, Z, in the wavefunctions to an effective nuclear charge (Z_{eff}), from +2 to a smaller value. The rationale for making this modification is that one electron partially shields the nuclear charge from the other electron, as shown in Figure 7.1.1 .

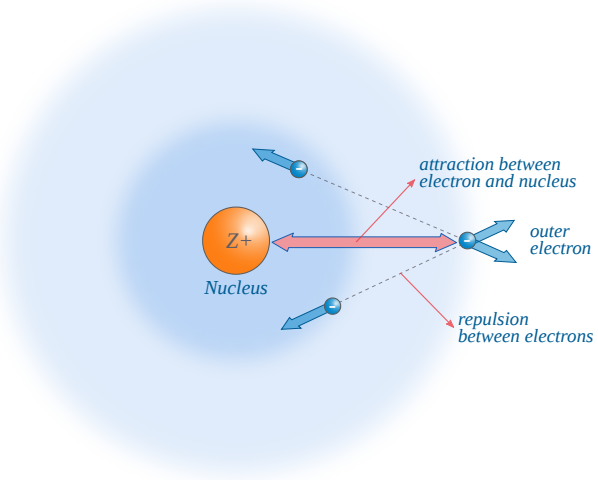


Figure 7.1.1 : Electron-electron shielding leading to a reduced effective nuclear charge. The attractive force of the nucleus on the outer electron is partially countered by the repulsive force between inner electrons and the outer electron. (CC BY-SA; Ümit Kaya via LibreTexts)

A region of negative charge density between one of the electrons and the +2 nucleus makes the potential energy between them more positive (decreases the attraction between them). We can effect this change mathematically by using $\zeta < 2$ in the wavefunction expression. If the shielding were complete, then Z_{eff} would equal 1. If there is no shielding, then $Z_{eff} = 2$. So a way to take into account the electron-electron interaction is by saying it produces a shielding effect. The shielding is not zero, and it is not complete, so the effective nuclear charge varies between one and two.

In general, a theory should be able to make predictions in advance of knowledge of the experimental result. Consequently, a principle and method for choosing the best value for Z_{eff} or any other adjustable parameter that is to be optimized in a calculation is needed. The Variational Principle provides the required criterion and method and says that the best value for any variable parameter in an approximate wavefunction is the value that gives the **lowest energy for the ground state**; i.e., the value that minimizes the energy. The variational method is the procedure that is used to find the lowest energy and the best values for the variable parameters.

A Better Approximation: The Variational Method

The variational method is one way of finding approximations to the lowest energy eigenstate or ground state, and some excited states. This allows calculating approximate wavefunctions and is the variational principle. The method consists in choosing a "trial wavefunction" depending on one or more parameters, and finding the values of these parameters for which the expectation value of the energy is the lowest possible. The wavefunction obtained by fixing the parameters to such values is then an approximation to the ground state wavefunction, and the expectation value of the energy in that state is an **upper bound** to the ground state energy.

The variational principle means that the expectation value for the binding energy obtained using an approximate wavefunction and the exact Hamiltonian operator will be higher than or equal to the true energy for the system. This idea is really powerful. When implemented, it permits us to find the best approximate wavefunction from a given wavefunction that contains one or more adjustable parameters, called a **trial wavefunction**. A mathematical statement of the variational principle is

$$E_{trial} \geq E_{true} \tag{7.1.7}$$

where

$$E_{trial} = \frac{\langle \psi_{trial} | \hat{H} | \psi_{trial} \rangle}{\langle \psi_{trial} | \psi_{trial} \rangle} \tag{7.1.8}$$

$$= \frac{\int \psi^*_{trial} \hat{H} \psi_{trial} d\tau}{\int \psi^*_{trial} \psi_{trial} d\tau} \tag{7.1.9}$$

Equation 7.1.7 is called the variational theorem and states that for a time-independent Hamiltonian operator, any trial wavefunction will have an variational energy (i.e., expectation value) that is **greater than or equal** to the true ground state wavefunction corresponding to the given Hamiltonian (Equation 7.1.7). Because of this, the variational energy is an upper bound to the true ground state energy of a given molecule. The general approach of this method consists in choosing a "trial wavefunction" depending on one or more parameters, and finding the values of these parameters for which the expectation value of the energy is the lowest possible (Figure 7.1.2).

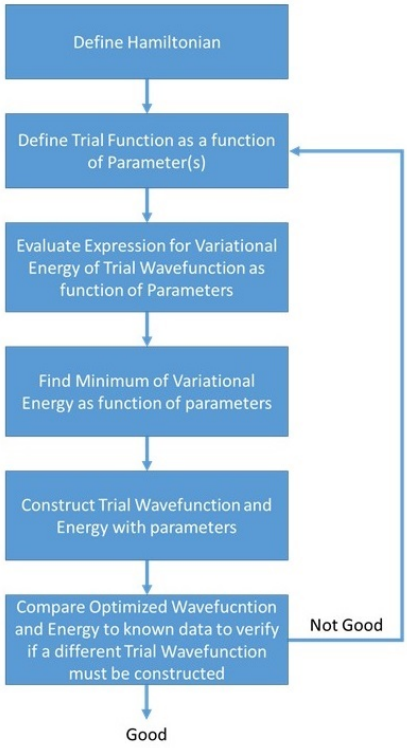


Figure 7.1.2 : Simplified algorithmic flowchart of the Variational Method approximation

> *The variational energy E_{trial} is only equal to the true energy E_{true} when the the corresponding trial wavefunction ψ_{trial} is equal to the true wavefunction ψ_{true}.*

Application to the Helium atom Ground State

Often the expectation values (numerator) and normalization integrals (denominator) in Equation 7.1.9 can be evaluated analytically. For the case of the He atom, let's consider the trial wavefunction as the product wavefunction given by Equation 7.1.10 (this is called the **orbital approximation**),

$$\psi(r_1, r_2) \approx \varphi(r_1)\varphi(r_2) \tag{7.1.10}$$

The adjustable or variable parameter in the trial wavefunction is the effective nuclear charge ζ, and the Hamiltonian is the complete form given below (Note: quantum calculations typically refer to effective nuclear charge as ζ rather than Z_{eff} as we used previously).

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\zeta e^2}{4\pi\epsilon_0 r_1} - \frac{\hbar^2}{2m}\nabla_2^2 - \frac{\zeta e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}} \tag{7.1.11}$$

the adjustable or variable parameter in the trial wavefunction is the effective nuclear charge ζ (would be equal to $\zeta = 2$ if fully unshielded), and the Hamiltonian is the complete form. When the expectation value for the trial energy (Equation 7.1.9) is evaluated for helium, the result is a variational energy that depends on the adjustable parameter, ζ.

$$E_{trial}(\zeta) = \frac{\mu e^4}{4\epsilon_0^2 h}\left(\zeta^2 - \frac{27}{8}\zeta\right) \tag{7.1.12}$$

This function is plotted in Figure 7.1.3 as a function of ζ. According to the variational principle (Equation 7.1.7), the minimum value of the energy on this graph is the best approximation of the true energy of the system, and the associated value of ζ is the best value for the adjustable parameter.

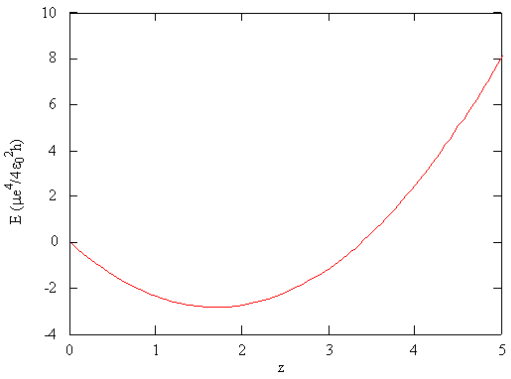


Figure 7.1.3 : Graph of trial energies for helium atom as a function of the adjustable parameter ζ in Equation 7.1.12 that represents the effective nuclear charge felt by the electrons.

Using the mathematical function for the energy of a system, the minimum energy with respect to the adjustable parameter can be found by taking the derivative of the energy with respect to that parameter, setting the resulting expression equal to zero, and solving for the parameter, in this case ζ. This is a standard method in calculus for finding maxima and minima.

? Exercise 7.1.1

Find the value for ζ that minimizes the helium binding energy for the product trial wavefunction in Equation 7.1.10 with the Hamiltonian in Equation 7.1.11. and compare the binding energy to the experimental value. What is the percent error in the calculated value?

Solution

The variational method requires following the workflow in Figure 7.1.2 .

- Step 1: Define the Hamiltonian - This is given by Equation 7.1.11.
- Step 2: Define the trial wavefunction as a function of at least one parameter - This is given by Equation 7.1.10.
- Step 3: Evaluate variational energy (E_{trial} integral (Equation 7.1.9) - This procedure was already above in Equation 7.1.12.
- Step 4: Minimize the variational energy as a function of the parameter(s) - Following the standard approach to find extrema in calculus, evaluate the derivative of E_{trial} with respect to ζ and set to zero:

$$\frac{dE_{trial}}{d\zeta} = \frac{\mu e^4}{4\epsilon_0^2 h}\left(2\zeta - \frac{27}{8}\right) = 0$$

then find solve for the roots of this polynomial

$$2\zeta - \frac{27}{8} = 0$$

or

$$\zeta = \frac{27}{16} \approx 1.6875$$

- Step 5-6: The question does not ask for the optimized wavefunction (Step 5) or to compare the result with the true value to evaluate the quality of the approximation (Step 6). We can skip these steps.

From Exercise 7.1.1 , the $\zeta = 1.6875$ and the approximate energy we calculate using this approximation method, E_{approx} = -77.483 eV. Table 7.1.1 show that a substantial improvement in the accuracy of the computed binding energy is obtained by using shielding to account for the electron-electron interaction. Including the effect of electron shielding in the wavefunction reduces the error in the binding energy to about 2%. This idea is very simple, elegant, and significant.

Table 7.1.1 : Comparison of the results of three approximation methods to experiment.

Method	He binding energy (eV)
"Ignorance is Bliss" Approximation (neglect repulsion between electrons)	-108.8
Variational method with variable effective charge	-77.483
Experimental	-79.0

The improvement we have seen in the total energy calculations using a variable parameter ζ indicates that an important contribution of electron-electron interaction or repulsion to the total binding energy arises from the fact that each electron shields the nuclear charge from the other electron. It is reasonable to assume the electrons are independent; i.e., that they move independently, but the shielding must be taken into account in order to fine-tune the wavefunctions. The inclusion of optimizable parameters in the wavefunction allows us to develop a clear physical image of the consequences of our variation calculation. Calculating energies correctly is important, and it is also important to be able to visualize electron densities for multi-electron systems. In the next two sections, we take a temporary break from our consideration of approximation methods in order to examine multi-electron wavefunctions more closely.

7.1: The Variational Method is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **7.1: The Variational Method Approximation** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0. Original source: https://web.archive.org/web/20200619182410/http://www.chemeddl.org.

7.2: Linear Variational Method and the Secular Determinant

Learning Objectives

- Understand how the variational method can be expanded to include trial wavefunctions that are a linear combination of functions with coefficients that are the parameters to be varied.
- To be able to construct secular equations to solve the minimization procedure intrinsic to the variational method approach.
- To map the secular equations into the secular determinant
- To understand how the Linear Combination of Atomic Orbital (LCAO) approximation is a specific application of the linear variational method.

A special type of variation widely used in the study of molecules is the so-called **linear variation function**, where the trial wavefunction is a linear combination of N linearly independent functions (often atomic orbitals) that not the eigenvalues of the Hamiltonian (since they are not known). For example

$$|\psi_{trial}\rangle = \sum_{j=1}^{N} a_j |\phi_j\rangle \tag{7.2.1}$$

and

$$\langle\psi_{trial}| = \sum_{j=1}^{N} a_j^* \langle\phi_j| \tag{7.2.2}$$

In these cases, one says that a 'linear variational' calculation is being performed.

Linear Variational Basis Functions

The set of functions $\{\phi_j\}$ are called the 'linear variational' basis functions and are nothing more than members of a set of functions that are convenient to deal with. However, they are typically not arbitrary and are usually selected to address specific properties of the system:

- to obey all of the boundary conditions that the exact state $|\psi_{trial}\rangle$ obeys,
- to be functions of the the same coordinates as $|\psi_{trial}\rangle$,
- to be of the same symmetry as $|\psi_{trial}\rangle$, and
- to be convenient to evaluate Hamiltonian terms elements $\langle\phi_i|H|\phi_j\rangle$.

Beyond these conditions, nothing other than effort can limit the selection and number of such basis functions in the expansions in Equations 7.2.1 and 7.2.2.

As discussed in Section 7.1, the variational energy for a generalized trial wavefunction is

$$E_{trial} = \frac{\langle\psi_{trial}|\hat{H}|\psi_{trial}\rangle}{\langle\psi_{trial}|\psi_{trial}\rangle} \tag{7.2.3}$$

Substituting Equations 7.2.1 and 7.2.2 into Equation 7.2.3 involves addressing the numerator and denominator individually. For the numerator, the integral can be expanded thusly:

$$\langle\psi_{trial}|H|\psi_{trial}\rangle = \sum_{i}^{N}\sum_{j}^{N} a_i^* a_j \langle\phi_i|H|\phi_j\rangle. \tag{7.2.4}$$

$$= \sum_{i,j}^{N,N} a_i^* a_j \langle\phi_i|H|\phi_j\rangle. \tag{7.2.5}$$

We can rewrite the following integral in Equation 7.2.5 as a function of the basis elements (not the trial wavefunction) as

$$H_{ij} = \langle\phi_i|H|\phi_j\rangle$$

So the numerator of the right side of Equation 7.2.3 becomes

$$\langle \psi_{trial}|H|\psi_{trial}\rangle = \sum_{i,j}^{N,N} a_i^* a_j H_{ij} \tag{7.2.6}$$

Similarly, the denominator of the right side of Equation 7.2.3 can be expanded

$$\langle \psi_{trial}|\psi_{trial}\rangle = \sum_{i,j}^{N,N} a_i^* a_j \langle \phi_i|\phi_j\rangle \tag{7.2.7}$$

We often simplify the integrals on the right side of Equation 7.2.7 as

$$S_{ij} = \langle \phi_i|\phi_j\rangle$$

where S_{ij} are **overlap integrals** between the different $\{\phi_j\}$ basis functions. Equation 7.2.7 is thus expressed as

$$\langle \psi_{trial}|\psi_{trial}\rangle = \sum_{i,j}^{N,N} a_i^* a_j S_{ij} \tag{7.2.8}$$

Orthonormality of Basis Functions

There is no explicit rule that the $\{\phi_j\}$ functions have to be **orthogonal** or **normalized** functions, although they often are selected that way for convenience. Therefore, *a priori*, S_{ij} does not have to be δ_{ij}.

Substituting Equations 7.2.6 and 7.2.8 into the variational energy formula (Equation 7.2.3) results in

$$E_{trial} = \frac{\sum_{i,j}^{N,N} a_i^* a_j H_{ij}}{\sum_{i,j}^{N,N} a_i^* a_j S_{ij}} \tag{7.2.9}$$

For such a trial wavefunction as Equation 7.2.1, the variational energy depends quadratically on the 'linear variational' a_j coefficients. These coefficients can be varied just like the parameters in the trial functions of Section 7.1 to find the optimized trial wavefunction ($|\psi_{trial}\rangle$) that approximates the true wavefunction ($|\psi\rangle$) that we cannot analytically solve for.

Minimizing the Variational Energy

The expression for variational energy (Equation 7.2.9) can be rearranged

$$E_{trial} \sum_{i,j}^{N,N} a_i^* a_j S_{ij} = \sum_{i,j}^{N,N} a_i^* a_j H_{ij} \tag{7.2.10}$$

The optimum coefficients are found by searching for minima in the variational energy landscape spanned by varying the $\{a_i\}$ coefficients (Figure 7.2.1).

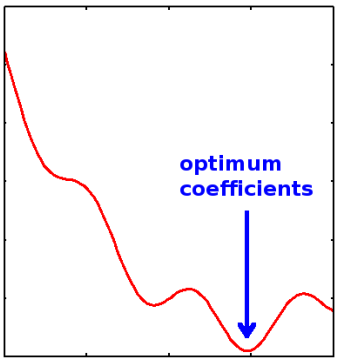


Figure 7.2.1 : Linear Variational Principle in action. Two (or more) wavefunctions are mixed by linear combination. The coefficients a_1, a_2 determine the weight each of them is given. The optimum coefficients are found by searching for minima in the potential landscape spanned by a_1 and a_2. (CC BY-SA-NC; Rudolf Winter; Aberystwyth University).

We want to minimize the energy with respect to the linear coefficients $\{a_i\}$, which requires that

$$\frac{\partial E_{trial}}{\partial a_i} = 0$$

for all i.

Differentiating both sides of Equation 7.2.10 for the kth coefficient gives,

$$\frac{\partial E_{trial}}{\partial a_k} \sum_{i,j}^{N,N} a_i^* a_j S_{ij} + E_{trial} \sum_i \sum_j \left[\frac{\partial a_i^*}{\partial a_k} a_j + \frac{\partial a_j}{\partial a_k} a_i^* \right] S_{ij} \tag{7.2.11}$$
$$= \sum_{i,j}^{N,N} \left[\frac{\partial a_i^*}{\partial a_k} a_j + \frac{\partial a_j}{\partial a_k} a_i^* \right] H_{ij}$$

Since the coefficients are independent

$$\frac{\partial a_i^*}{\partial a_k} = \delta_{ik}$$

and

$$S_{ij} = S_{ji}$$

and also since the Hamiltonian is a Hermitian Operator (see below)

$$H_{ij} = H_{ji}$$

then Equation 7.2.11 simplifies to

$$\frac{\partial E_{trial}}{\partial a_k} \sum_i \sum_j a_i^* a_j S_{ij} + 2E_{trial} \sum_i a_i S_{ik} = 2 \sum_i a_i H_{ik} \tag{7.2.12}$$

At the minimum variational energy, when

$$\frac{\partial E_{trial}}{\partial a_k} = 0$$

then Equation 7.2.12 gives

$$\sum_i^N a_i (H_{ik} - E_{trial} S_{ik}) = 0 \tag{7.2.13}$$

for all k. The equations in 7.2.13 are called the **Secular Equations**.

Hermitian Operators

Hermitian operators are operators that satisfy the general formula

$$\langle \phi_i | \hat{A} | \phi_j \rangle = \langle \phi_j | \hat{A} | \phi_i \rangle \tag{7.2.14}$$

If that condition is met, then $\hat{A}$ is a Hermitian operator. For any operator that generates a real eigenvalue (e.g., observables), then that operator is Hermitian. The Hamiltonian $\hat{H}$ meets the condition of a Hermitian operator. Equation 7.2.14 can be rewriten as

$$A_{ij} = A^*_{ji}$$

where

$$A_{ij} = \langle \phi_i | \hat{A} | \phi_j \rangle$$

and

$$A_{ji} = \langle \phi_j | \hat{A} | \phi_i \rangle$$

Therefore, when applied to the Hamiltonian operator

$$H^*_{ij} = H_{ji}.$$

If the functions $\{|\phi_j\rangle\}$ are orthonormal, then the overlap matrix S reduces to the unit matrix (one on the diagonal and zero every where else) and the Secular Equations in Equation 7.2.13 reduces to the more familiar Eigenvalue form:

$$\sum_i^N H_{ij} a_j = E_{trial} a_i. \tag{7.2.15}$$

Hence, the secular equation, in either form, have as many eigenvalues E_i and eigenvectors $\{C_{ij}\}$ as the dimension of the H_{ij} matrix as the functions in $|\psi_{trail}\rangle$ (Example 7.2.1). It can also be shown that between successive pairs of the eigenvalues obtained by solving the secular problem at least one exact eigenvalue must occur (i.e.,$E_{i+1} > E_{exact} > E_i$, for all i). This observation is referred to as '*the bracketing theorem*'.

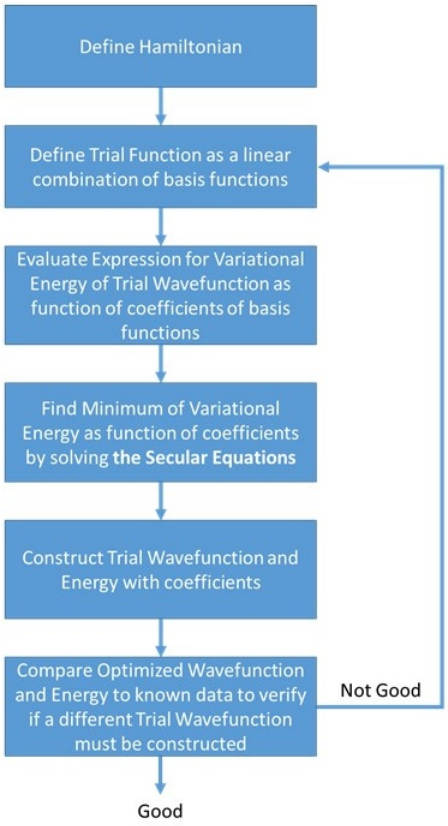


Figure 7.2.2 : Simplified algorithmic flowchart of the linear Variational Method approximation. Compared to the similar algorithmic flowchart for the "normal" Variational Method approximation.

Variational methods, in particular the linear variational method, are the most widely used approximation techniques in quantum chemistry. To implement such a method one needs to know the Hamiltonian H whose energy levels are sought and one needs to construct a trial wavefunction in which some 'flexibility' exists (e.g., as in the linear variational method where the a_j coefficients can be varied). This tool will be used to develop several of the most commonly used and powerful molecular orbital methods in chemistry.

The Secular Determinant

From the secular equations with an orthonormal functions (Equation 7.2.15), we have k simultaneous secular equations in k unknowns. These equations can also be written in matrix notation, and for a non-trivial solution (i.e. $c_i \neq 0$ for all i), the determinant of the secular matrix must be equal to zero.

$$|H_{ik} - ES_{ik}| = 0 \tag{7.2.16}$$

Properties of Determinants

- The determinant is a **real** number, it is **not** a matrix.
- The determinant can be a negative number.
- It is **not** associated with absolute value at all except that they both use vertical lines.
- The determinant **only** exists for square matrices (2×2, 3×3, ..., $n \times n$). The determinant of a 1×1 matrix is that single value in the determinant.
- The inverse of a matrix will exist only if the determinant is not zero.

Expanding Determinants

The determinant can be evaluated using an expansion method involving minors and cofactors. Before we can use them, we need to define them. It is the product of the elements on the main diagonal minus the product of the elements off the main diagonal. In the case of a 2×2 matrix, the specific formula for the determinant is

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Similarly, suppose we have a 3×3 matrix A, and we want the specific formula for its determinant $|A|$:

$$\begin{aligned} |A| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ &= aei + bfg + cdh - ceg - bdi - afh. \end{aligned}$$

To solve Equation 7.2.16, the determinate should be expanded and then set to zero. That generates a polynomial (called a **characteristic equation**) that can be directly solved with linear algebra methods or numerically.

Example 7.2.1 : A Simple Two Component Basis Set

If $|\psi_{trial}\rangle$ is a linear combination of two functions. In math terms,

$$|\psi_{trial}\rangle = \sum_{n=1}^{N=2} a_n |f_n\rangle = a_1|\phi_1\rangle + a_2|\phi_2\rangle$$

then the secular determinant (Equation 7.2.16), in matrix formulation would look like this

$$\begin{vmatrix} H_{11} - E_{trial}S_{11} & H_{12} - E_{trial}S_{12} \\ H_{12} - E_{trial}S_{12} & H_{22} - E_{trial}S_{22} \end{vmatrix} = 0$$

Solution

Solving the secular equations is done by finding E_{trial} and putting the value into the expansion of the secular determinant

$$a_1^2 H_{11} + 2a_1 a_2 H_{12} + a_2^2 H_{22} = 0$$

and

$$a_1(H_{12} - E_{trial}S_{12}) + a_2(H_{22} - E_{trial}S_{22}) = 0$$

Equation 7.2.16 can be solved to obtain the energies E. When arranged in order of increasing energy, these provide approximations to the energies of the first k states (each having an energy higher than the true energy of the state by virtue of the variation theorem). To find the energies of a larger number of states we simply use a **greater number** of basis functions $\{\phi_i\}$ in the trial wavefunction (Example 7.2.1). To obtain the approximate wavefunction for a particular state, we substitute the appropriate energy into the secular equations and solve for the coefficients a_i.

Using this method it is possible to find all the coefficients $a_1 \ldots a_k$ in terms of one coefficient; normalizing the wavefunction provides the absolute values for the coefficients.

Example 7.2.2 : Linear Combination of Atomic Orbitals (LCAO) Approximation

Trial wavefunctions that consist of linear combinations of simple functions

$$|\psi(r)\rangle = \sum_i a_i |\phi_i(r)\rangle$$

form the basis of the Linear Combination of Atomic Orbitals (LCAO) method introduced by Lennard and Jones and others to compute the energies and wavefunctions of atoms and molecules. The functions $\{|\phi_i\rangle\}$ are selected so that matrix elements can be evaluated analytically. Two basis sets of atomic orbitals functions can be used: Slater type and Gaussian type:

Slater orbitals using Hydrogen-like wavefunctions

$$|\phi_i\rangle = Y_l^m(\theta,\phi)e^{-\alpha r}$$

and Gaussian orbitals of the form

$$|\phi_i\rangle = Y_l^m(\theta,\phi)e^{-\alpha r^2}$$

are the most widely used forms, where $Y_l^m(\theta,\phi)$ are the spherical harmonics that represent the angular part of the atomic orbitals. Gaussian orbitals form the basis of many quantum chemistry computer codes.

Dec 23, 2018, 11:08 PM

Because Slater orbitals give exact results for Hydrogen, we will use Gaussian orbitals to test the LCAO method on Hydrogen, following S.F. Boys, Proc. Roy. Soc. A 200, 542 (1950) and W.R. Ditchfield, W.J. Hehre and J.A. Pople, J. Chem. Phys. Rev. 52, 5001 (1970) with the basis set. Because products of Gaussians are also Gaussian, the required matrix elements are easily computed.

The linear variational method is used extensively in molecular orbitals of molecules and further examples will be postponed until that discussion in Chapters 9.

This page titled 7.2: Linear Variational Method and the Secular Determinant is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Claire Vallance via source content that was edited to the style and standards of the LibreTexts platform.

- **7.2: Linear Variational Method and the Secular Determinant** by Claire Vallance is licensed CC BY-NC-SA 4.0. Original source: http://vallance.chem.ox.ac.uk/pdfs/SymmetryLectureNotes.pdf.

7.3: Trial Functions Can Be Linear Combinations of Functions with Variational Parameters

Learning Objectives

- Demonstrate that variational problems can include changing parameters within the elements (normal variational method) and changing coefficients of a basis set (linear variational method)

An alternative approach to the general problem of introducing variational parameters into wavefunctions is the construction of a wavefunction as a linear combination of other functions each with one or multiple parameters that can be varied

For hydrogen, the radial function decays, or decreases in amplitude, exponentially as the distance from the nucleus increases. For helium and other multi-electron atoms, the radial dependence of the total probability density does not fall off as a simple exponential with increasing distance from the nucleus as it does for hydrogen. More complex single-electron functions therefore are needed in order to model the effects of electron-electron interactions on the total radial distribution function. One way to obtain more appropriate single-electron functions is to use a sum of exponential functions in place of the hydrogenic spin-orbitals.

An example of such a wavefunction created from a sum or linear combination of exponential functions is written as

$$\varphi_{1s}(r_1) = \sum_j c_j e^{-\zeta_j r_j / a_o} \tag{7.3.1}$$

The linear combination permits weighing of the different exponentials through the adjustable coefficients (c_j) for each term in the sum. Each exponential term has a different rate of decay through the zeta-parameter ζ_j. The exponential functions in Equation 7.3.1 are called **basis functions**. Basis functions are the functions used in linear combinations to produce the single-electron orbitals that in turn combine to create the product multi-electron wavefunctions. Originally the most popular basis functions used were the STO's, but today STO's are not used in most quantum chemistry calculations. However, they are often the functions to which more computationally efficient basis functions are fitted.

Physically, the ζ_j parameters account for the effective nuclear charge (often denoted with Z_{eff}). The use of several zeta values in the linear combination essentially allows the effective nuclear charge to vary with the distance of an electron from the nucleus. This variation makes sense physically. When an electron is close to the nucleus, the effective nuclear charge should be close to the actual nuclear charge. When the electron is far from the nucleus, the effective nuclear charge should be much smaller. See Slater's rules for a rule-of-thumb approach to evaluate Z_{eff} values.

A term in Equation 7.3.1 with a small ζ will decay slowly with distance from the nucleus. A term with a large ζ will decay rapidly with distance and not contribute at large distances. The need for such a linear combination of exponentials is a consequence of the electron-electron repulsion and its effect of screening the nucleus for each electron due to the presence of the other electrons.

Exercise 7.3.1

Make plots of φ in Equation 7.3.1 using three equally weighted terms with ζ = 1.0, 2.0, and 5.0. Also plot each term separately.

Computational procedures in which an exponential parameter like ζ is varied are more precisely called the Nonlinear Variational Method because the variational parameter is part of the wavefunction and the change in the function and energy caused by a change in the parameter is not linear. The optimum values for the zeta parameters in any particular calculation are determined by doing a variational calculation for each orbital to minimize the ground-state energy. When this calculation involves a nonlinear variational calculation for the zetas, it requires a large amount of computer time. The use of the variational method to find values for the coefficients, $\{c_j\}$, in the linear combination given by Equation 7.3.1 above is called the Linear Variational Method because the single-electron function whose energy is to be minimized (in this case φ_{1s}) depends linearly on the coefficients. Although the idea is the same, it usually is much easier to implement the linear variational method in practice.

Nonlinear variational calculations are extremely costly in terms of computer time because each time a zeta parameter is changed, all of the integrals need to be recalculated. In the linear variation, where only the coefficients in a linear combination are varied, the

basis functions and the integrals do not change. Consequently, an optimum set of zeta parameters were chosen from variational calculations on many small multi-electron systems, and these values, which are given in Table 7.3.1 , generally can be used in the STOs for other and larger systems.

Table 7.3.1 : Orbital Exponents for Slater Orbitals

Atom	ζ_{1s}	$\zeta_{2s,2p}$
H	1.24	-
He	1.69	-
Li	2.69	0.80
Be	3.68	1.15
B	4.68	1.50
C	5.67	1.72
N	6.67	1.95
O	7.66	2.25
F	8.56	2.55
Ne	9.64	2.88

? Exercise 7.3.2

Compare the value ζ_{1s} = 1.24 in Table 7.3.1 for hydrogen with the value you obtained in Exercise 7.3.1 . and comment on possible reasons for any difference. Why are the zeta values larger for 1s than for 2s and 2p orbitals? Why do the ζ_{1s} values increase by essentially one unit for each element from He to Ne while the increase for the $\zeta_{2s,2p}$ values is much smaller?

The discussion above gives us some new ideas about how to write flexible, useful single-electron wavefunctions that can be used to construct multi-electron wavefunctions for variational calculations. Single-electron functions built from the basis function approach are flexible because they have several adjustable parameters, and useful because the adjustable parameters still have clear physical interpretations. Such functions will be needed in the Hartree-Fock method discussed elsewhere.

7.3: Trial Functions Can Be Linear Combinations of Functions with Variational Parameters is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **7.3: Trial Functions Can Be Linear Combinations of Functions That Also Contain Variational Parameters** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

7.4: Perturbation Theory Expresses the Solutions in Terms of Solved Problems

Learning Objectives

- Use perturbation theory to approximate the energies of systems as a series of perturbation of a solved system.
- Use perturbation theory to approximate the wavefunctions of systems as a series of perturbation of a solved system.

It is easier to compute the changes in the energy levels and wavefunctions with a scheme of successive corrections to the zero-field values. This method, termed **perturbation theory**, is the single most important method of solving problems in quantum mechanics and is widely used in atomic physics, condensed matter and particle physics. Perturbation theory is another approach to finding approximate solutions to a problem, by starting from the exact solution of a related, simpler problem. A critical feature of the technique is a middle step that breaks the problem into "solvable" and "perturbation" parts. Perturbation theory is applicable if the problem at hand cannot be solved exactly, but can be formulated by adding a "small" term to the mathematical description of the exactly solvable problem.

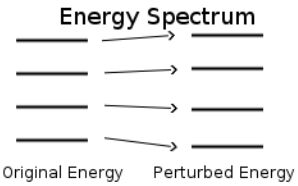


Figure 7.4.1 : Perturbed Energy Spectrum. (CC BY-SA 2.0, Frontier via Wikipedia)

We begin with a Hamiltonian $\hat{H}^0$ having known eigenstates and eigenenergies:

$$\hat{H}^o|n^o\rangle = E_n^o|n^o\rangle \tag{7.4.1}$$

The task is to find how these eigenstates and eigenenergies change if a small term H^1 (an external field, for example) is added to the Hamiltonian, so:

$$(\hat{H}^0 + \hat{H}^1)|n\rangle = E_n|n\rangle \tag{7.4.2}$$

That is to say, on switching on $\hat{H}^1$ changes the wavefunctions:

$$\underbrace{|n^o\rangle}_{\text{unperturbed}} \Rightarrow \underbrace{|n\rangle}_{\text{Perturbed}} \tag{7.4.3}$$

and energies (Figure 7.4.1):

$$\underbrace{E_n^o}_{\text{unperturbed}} \Rightarrow \underbrace{E_n}_{\text{Perturbed}} \tag{7.4.4}$$

> *The basic assumption in perturbation theory is that H^1 is sufficiently small that the leading corrections are the same order of magnitude as H^1 itself, and the true energies can be better and better approximated by a successive series of corrections, each of order H^1/H^o compared with the previous one.*

The strategy is to **expand** the true wavefunction and corresponding eigenenergy as series in $\hat{H}^1/\hat{H}^o$. These series are then fed into Equation 7.4.2, and terms of the same order of magnitude in $\hat{H}^1/\hat{H}^o$ on the two sides are set equal. The equations thus generated are solved one by one to give progressively more accurate results.

To make it easier to identify terms of the same order in $\hat{H}^1/\hat{H}^o$ on the two sides of the equation, it is convenient to introduce a dimensionless parameter λ which always goes with $\hat{H}^1$, and then expand both eigenstates and eigenenergies as power series in λ,

$$|n\rangle = \sum_i^m \lambda^i |n^i\rangle \tag{7.4.5}$$

$$E_n = \sum_{i=0}^m \lambda^i E_n^i \tag{7.4.6}$$

where m is how many terms in the expansion we are considering. The ket $|n^i\rangle$ is multiplied by λ^i and is therefore of order $(H^1/H^o)^i$.

> *λ is purely a bookkeeping device: we will set it equal to 1 when we are through! It's just there to keep track of the orders of magnitudes of the various terms.*

For example, in first order perturbation theory, Equations 7.4.5 are truncated at $m = 1$ (and setting $\lambda = 1$):

$$|n\rangle \approx |n^o\rangle + |n^1\rangle \tag{7.4.7}$$

$$E_n \approx E_n^o + E_n^1 \tag{7.4.8}$$

However, let's consider the general case for now. Adding the full expansions for the eigenstate (Equation 7.4.5) and energies (Equation 7.4.6) into the Schrödinger equation for the perturbation Equation 7.4.2 in

$$(\hat{H}^o + \lambda \hat{H}^1)|n\rangle = E_n|n\rangle \tag{7.4.9}$$

we have

$$(\hat{H}^o + \lambda \hat{H}^1)\left(\sum_{i=0}^m \lambda^i |n^i\rangle\right) = \left(\sum_{i=0}^m \lambda^i E_n^i\right)\left(\sum_{i=0}^m \lambda^i |n^i\rangle\right) \tag{7.4.10}$$

We're now ready to match the two sides term by term in powers of λ. Note that the zeroth-order term, of course, just gives back the unperturbed Schrödinger Equation (Equation 7.4.1).

Let's look at Equation 7.4.10 with the first few terms of the expansion:

$$(\hat{H}^o + \lambda \hat{H}^1)\left(|n^o\rangle + \lambda|n^1\rangle\right) = \left(E_n^0 + \lambda E_n^1\right)\left(|n^o\rangle + \lambda|n^1\rangle\right) \tag{7.4.11}$$

$$\begin{aligned} \hat{H}^o|n^o\rangle + \lambda\hat{H}^1|n^o\rangle + \lambda H^o|n^1\rangle + \lambda^2\hat{H}^1|n^1\rangle = {} & E_n^0|n^o\rangle + \lambda E_n^1|n^o\rangle \\ & + \lambda E_n^0|n^1\rangle + \lambda^2 E_n^1|n^1\rangle \end{aligned} \tag{7.4.12}$$

Collecting terms in order of λ and coloring to indicate different orders

$$\begin{aligned} & \underset{\text{zero order}}{\hat{H}^o|n^o\rangle} \\ & + \underset{\text{1st order}}{\lambda(\hat{H}^1|n^o\rangle + \hat{H}^o|n^1\rangle)} + \underset{\text{2nd order}}{\lambda^2\hat{H}^1|n^1\rangle} = \underset{\text{zero order}}{E_n^0|n^o\rangle} + \underset{\text{1st order}}{\lambda(E_n^1|n^o\rangle + E_n^0|n^1\rangle)} + \underset{\text{2nd order}}{\lambda^2 E_n^1|n^1\rangle} \end{aligned} \tag{7.4.13}$$

If we expanded Equation 7.4.10 further we could express the energies and wavefunctions in higher order components.

Zero-Order Terms ($\lambda = 0$)

Collecting the zero order terms in the expansion (black terms in Equation 7.4.13) results in just the Schrödinger Equation for the unperturbed system

$$\hat{H}^o|n^o\rangle = E_n^o|n^o\rangle \tag{7.4.14}$$

First-Order Expression of Energy ($\lambda = 1$)

The summations in Equations 7.4.5, 7.4.6, and 7.4.10 can be truncated at any order of λ. For example, the first order perturbation theory has the truncation at $\lambda = 1$. Matching the terms that linear in λ (red terms in Equation 7.4.13) and setting $\lambda = 1$ on both sides of Equation 7.4.13:

$$\hat{H}^o|n^1\rangle + \hat{H}^1|n^o\rangle = E_n^o|n^1\rangle + E_n^1|n^o\rangle \tag{7.4.15}$$

Equation 7.4.15 is the key to finding the first-order change in energy E_n^1. Taking the **inner product** of both sides with $\langle n^o|$:

$$\langle n^o|\hat{H}^o|n^1\rangle + \langle n^o|\hat{H}^1|n^o\rangle = \langle n^o|E_n^o|n^1\rangle + \langle n^o|E_n^1|n^o\rangle \tag{7.4.16}$$

since operating the zero-order Hamiltonian on the bra wavefunction (this is just the Schrödinger equation; Equation 7.4.14) is

$$\langle n^o|\hat{H}^o = \langle n^o|E_n^o \tag{7.4.17}$$

and via the orthonormality of the unperturbed $|n^o\rangle$ wavefunctions both

$$\langle n^o|n^o\rangle = 1 \tag{7.4.18}$$

and Equation 7.4.8 can be simplified

$$\cancel{E_n^o\langle n^o|n^1\rangle} + \langle n^o|H^1|n^o\rangle = \cancel{E_n^o\langle n^o|n^1\rangle} + E_n^1\cancel{\langle n^o|n^o\rangle^1} \tag{7.4.19}$$

since the unperturbed set of eigenstates are orthogonal (Equation 7.4.18) and we can cancel the other term on each side of the equation, we find that

$$E_n^1 = \langle n^o|\hat{H}^1|n^o\rangle \tag{7.4.20}$$

> *The first-order change in the energy of a state resulting from adding a perturbing term $\hat{H}^1$ to the Hamiltonian is just the **expectation value** of $\hat{H}^1$ in the **unperturbed** wavefunctions.*

That is, the first order energies (Equation 7.4.15) are given by

$$E_n \approx E_n^o + E_n^1 \tag{7.4.21}$$

$$\approx \underbrace{E_n^o + \langle n^o|H^1|n^o\rangle}_{\text{First Order Perturbation}} \tag{7.4.22}$$

7.4.1A ✔ Example : A Perturbed Particle in a Box

Estimate the energy of the ground-state and first excited-state wavefunction within first-order perturbation theory of a system with the following potential energy

$$V(x) = \begin{cases} V_o & 0 \le x \le L \\ \infty & x < 0 \text{ and } x > L \end{cases}$$

Solution

The first step in any perturbation problem is to write the Hamiltonian in terms of a unperturbed component that the solutions (both eigenstates and energy) are known and a perturbation component (Equation 7.4.2). For this system, the unperturbed Hamilonian and solutions is the particle in an infiinitely high box and the perturbation is a shift of the potential within the box by V_o.

$$\hat{H}^1 = V_o$$

Using Equation 7.4.20 for the first-order term in the energy of the ground-state

$$E_n^1 = \langle n^o|H^1|n^o\rangle$$

with the wavefunctions known from the particle in the box problem

$$|n^o\rangle = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}x\right)$$

At this stage we can do two problems independently (i.e., the ground-state with $|1\rangle$ and the first excited-state $|2\rangle$). However, in this case, the first-order perturbation to any particle-in-the-box state can be easily derived.

$$E_n^1 = \int_0^L \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}x\right) V_o \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}x\right) dx$$

or better yet, instead of evaluating this integrals we can simplify the expression

$$E_n^1 = \langle n^o|H^1|n^o\rangle = \langle n^o|V_o|n^o\rangle = V_o\langle n^o|n^o\rangle = V_o$$

so via Equation 7.4.22, the energy of each perturbed eigenstate is

$$\begin{aligned} E_n &\approx E_n^o + E_n^1 \\ &\approx \frac{h^2}{8mL^2}n^2 + V_o \end{aligned}$$

While this is the first order perturbation to the energy, it is also the exact value.

7.4.1B ✔ Example : An Even More Perturbed Particle in a Box

Estimate the energy of the ground-state wavefunction within first-order perturbation theory of a system with the following potential energy

$$V(x) = \begin{cases} V_o & 0 \le x \le L/2 \\ \infty & x < 0 \; and \; x > L \end{cases}$$

Solution

As with Example 7.4.1 , we recognize that unperturbed component of the problem (Equation 7.4.2) is the particle in an infinitely high well. For this system, the unperturbed Hamiltonian and solution is the particle in an infinitely high box and the perturbation is a shift of the potential within half a box by V_o. This is essentially a step function.

Using Equation 7.4.20 for the first-order term in the energy of the any state

$$\begin{aligned} E_n^1 &= \langle n^o|H^1|n^o\rangle \\ &= \int_0^{L/2}\sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}x\right)V_o\sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}x\right)dx + \int_{L/2}^{L}\sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}x\right)0\sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}x\right)dx \end{aligned}$$

The second integral is zero and the first integral is simplified to

$$E_n^1 = \frac{2}{L}\int_0^{L/2} V_o \sin^2\left(\frac{n\pi}{L}x\right)dx$$

this is evaluated to

$$\begin{aligned} E_n^1 &= \frac{2V_o}{L}\left[\frac{-1}{2\frac{\pi n}{a}}\cos\left(\frac{n\pi}{L}x\right)\sin\left(\frac{n\pi}{L}x\right) + \frac{x}{2}\right]_0^{L/2} \\ &= \frac{2V_o}{\cancel{L}}\frac{\cancel{L}}{4} = \frac{V_o}{2} \end{aligned}$$

The energy of each perturbed eigenstate, via Equation 7.4.22, is

$$\begin{aligned} E_n &\approx E_n^o + \frac{V_o}{2} \\ &\approx \frac{h^2}{8mL^2}n^2 + \frac{V_o}{2} \end{aligned}$$

First-Order Expression of Wavefunction ($\lambda = 1$)

The general expression for the first-order change in the *wavefunction* is found by taking the inner product of the first-order expansion (Equation 7.4.15) with the bra $\langle m^o|$ with $m \neq n$,

$$\langle m^o|H^o|n^1\rangle + \langle m^o|H^1|n^o\rangle = \langle m^o|E_n^o|n^1\rangle + \langle m^o|E_n^1|n^o\rangle \tag{7.4.23}$$

Last term on right side of Equation 7.4.23

The last integral on the right hand side of Equation 7.4.23 is zero, since $m \neq n$ so

$$\langle m^o|E_n^1|n^o\rangle = E_n^1\langle m^o|n^o\rangle \tag{7.4.24}$$

and

$$\langle m^o|n^0\rangle = 0 \tag{7.4.25}$$

First term on right side of Equation 7.4.23

The first integral is more complicated and can be expanded back into the H^o

$$E_m^o\langle m^o|n^1\rangle = \langle m^o|E_m^o|n^1\rangle = \langle m^o|H^o|n^1\rangle \tag{7.4.26}$$

since

$$\langle m^o|H^o = \langle m^o|E_m^o \tag{7.4.27}$$

so

$$\langle m^o|n^1\rangle = \frac{\langle m^o|H^1|n^o\rangle}{E_n^o - E_m^o} \tag{7.4.28}$$

and therefore the wavefunction corrected to first order is:

$$|n\rangle \approx |n^o\rangle + |n^1\rangle \tag{7.4.29}$$

$$\approx |n^o\rangle + \underbrace{\sum_{m\neq n} \frac{|m^o\rangle\langle m^o|H^1|n^o\rangle}{E_n^o - E_m^o}}_{\text{First Order Perturbation Theory}} \tag{7.4.30}$$

Equation 7.4.30 is essentially is an expansion of the unknown wavefunction correction as a **linear combination** of known unperturbed wavefunctions 7.4.32:

$$|n\rangle \approx |n^o\rangle + |n^1\rangle \tag{7.4.31}$$

$$\approx |n^o\rangle + \sum_{m\neq n} c_{m,n}|m^o\rangle \tag{7.4.32}$$

with the expansion coefficients determined by

$$c_{m,n} = \frac{\langle m^o|H^1|n^o\rangle}{E_n^o - E_m^o} \tag{7.4.33}$$

This is justified since the set of original zero-order wavefunctions forms a **complete basis set** that can describe any function.

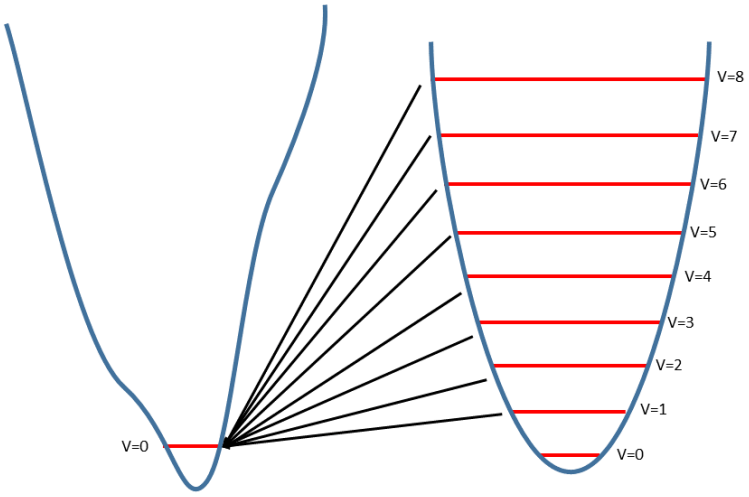


Figure 7.4.2 : The first order perturbation of the ground-state wavefunction for a perturbed (left potential) can be expressed as a linear combination of all excited-state wavefunctions of the unperturbed potential (Equation 7.4.32), shown as a harmonic oscillator in this example (right potential). Note that the ground-state harmonic oscillator wavefunction is not part of this expression and all wavefunctions need to be included in the expression, not just the first eight wavefunctions shown here. (CC BY 4.0; Delmar Larsen)

Calculating the first order perturbation to the wavefunctions is in general an infinite sum of off diagonal matrix elements of $\boldsymbol{H}^1$ (Figure 7.4.2).

- However, the denominator argues that terms in this sum will be weighted by states that are of **comparable energy**. That means in principle, these sum can be truncated easily based off of some criterion.
- Another point to consider is that many of these matrix elements will equal zero depending on the symmetry of the $\{|n^o\rangle\}$ basis and $\boldsymbol{H}^1$ (e.g., some $\langle m^o|\boldsymbol{H}^1|n^o\rangle$ integrals in Equation 7.4.30 could be zero due to the integrand having an odd symmetry; see Example 7.4.3).

> *The denominators in Equation 7.4.30 argues that terms in this sum will be preferentially dictated by states that are of **comparable energy**. That is, eigenstates that have energies significantly greater or lower than the unperturbed eigenstate will weakly contribute to the perturbed wavefunction.*

Example 7.4.2 : A Harmonic Oscillator with a Cubic Perturbation

Estimate the energy of the ground-state wavefunction associated with the Hamiltonian using perturbation theory

$$\hat{H} = \frac{-\hbar}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2 + \epsilon x^3$$

Solution

The first step in a perturbation theory problem is to identify the reference system with the known eigenstates and energies. For this example, this is clearly the harmonic oscillator model.

Energy

The first steps in flowchart for applying perturbation theory (Figure 7.4.1) is to separate the Hamiltonian of the difficult (or unsolvable) problem into a solvable one with a perturbation. For this case, we can rewrite the Hamiltonian as

$$\hat{H}^o + \hat{H}^1$$

where

- $\hat{H}^o$ is the Hamitonian for the standard Harmonic Oscillator with **known** eigenstates and eigenenergies

$$\hat{H}^{(0)} = \frac{-\hbar}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2$$

- $\hat{H}^1$ is the pertubtiation

$$\hat{H}^1 = \epsilon x^3$$

The first order perturbation is given by Equation 7.4.20, which for this problem is

$$E_n^1 = \langle n^o | \epsilon x^3 | n^o \rangle$$

Notice that the integrand has an odd symmetry (i.e., $f(x) = -f(-x)$) with the perturbation Hamiltonian being odd and the ground state harmonic oscillator wavefunctions being even. So

$$E_n^1 = 0$$

This means to first order pertubation theory, this cubic terms does not alter the ground state energy (via Equation 7.4.22). However, this is not the case if second-order perturbation theory were used, which is more accurate (not shown).

Wavefunction

Calculating the first order perturbation to the wavefunctions (Equation 7.4.30) is more difficult than energy since multiple integrals must be evaluated (an infinite number if symmetry arguments are not applicable). The harmonic oscillator wavefunctions are often written in terms of Q, the unscaled displacement coordinate:

$$|\Psi_v(x)\rangle = N_v'' H_v(\sqrt{\alpha}Q) e^{-\alpha Q^2/2}$$

with α

$$\alpha = 1/\sqrt{\beta} = \sqrt{\frac{k\mu}{\hbar^2}}$$

and

$$N_v'' = \sqrt{\frac{1}{2^v v!}} \left(\frac{\alpha}{\pi}\right)^{1/4}$$

Let's consider only the first six wavefunctions that use these Hermite polynomials $H_v(x)$:

- $H_0 = 1$
- $H_1 = 2x$
- $H_2 = -2 + 4x^2$
- $H_3 = -12x + 8x^3$
- $H_4 = 12 - 48x^2 + 16x^4$
- $H_5 = 120x - 160x^3 + 32x^5$

The first order perturbation to the ground-state wavefunction (Equation 7.4.30)

$$|0^1\rangle = \sum_{m \neq 0}^{5} \frac{|m^o\rangle\langle m^o|H^1|0^o\rangle}{E_0^o - E_m^o} \tag{7.4.34}$$

given these truncated wavefunctions (we should technically use the infinite sum) and that we are considering only the ground state with $n = 0$:

$$|0^1\rangle = \frac{\langle 1^o|H^1|0^o\rangle}{E_0^o - E_1^o}|1^o\rangle + \frac{\langle 2^o|H^1|0^o\rangle}{E_0^o - E_2^o}|2^o\rangle + \frac{\langle 3^o|H^1|0^o\rangle}{E_0^o - E_3^o}|3^o\rangle + \frac{\langle 4^o|H^1|0^o\rangle}{E_0^o - E_4^o}|4^o\rangle + \frac{\langle 5^o|H^1|0^o\rangle}{E_0^o - E_5^o}|5^o\rangle$$

We can use symmetry of the perturbation and unperturbed wavefunctions to solve the integrals above. We know that the unperturbed harmonic oscillator wavefunctions $\{|n^0\}\rangle$ alternate between even (when v is even) and odd (when v is odd) as shown previously. Since the perturbation is an odd function, only when $m = 2k + 1$ with $k = 1, 2, 3$ would these integrals be non-zero (i.e., for $m = 1, 3, 5, ...$).

So of the original five unperturbed wavefunctions, only $|m=1\rangle$, $|m=3\rangle$, and $|m=5\rangle$ mix to make the first-order perturbed ground-state wavefunction so

$$|0^1\rangle = \frac{\langle 1^o|H^1|0^o\rangle}{E_0^o - E_1^o}|1^o\rangle + \frac{\langle 3^o|H^1|0^o\rangle}{E_0^o - E_3^o}|3^o\rangle + \frac{\langle 5^o|H^1|0^o\rangle}{E_0^o - E_5^o}|5^o\rangle$$

At this stage, the integrals have to be manually calculated using the defined wavefuctions above, which is left as an exercise. Notice that each unperturbed wavefunction that can "mix" to generate the perturbed wavefunction will have a reciprocally decreasing contribution (w.r.t. energy) due to the growing denominator in Equation 7.4.34.

? Exercise 7.4.3 : Harmonic Oscillator with a Quartic Perturbation

Use perturbation theory to estimate the energy of the ground-state wavefunction associated with this Hamiltonian

$$\hat{H} = \frac{-\hbar}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2 + \gamma x^4.$$

Second-Order Terms ($\lambda = 2$)

There are higher energy terms in the expansion of Equation 7.4.5 (e.g., the blue terms in Equation 7.4.13), but are not discussed further here other than noting the whole perturbation process is an infinite series of corrections that ideally converge to the correct answer. It is truncating this series as a finite number of steps that is the approximation. The general approach to perturbation theory applications is giving in the flowchart in Figure 7.4.1 .

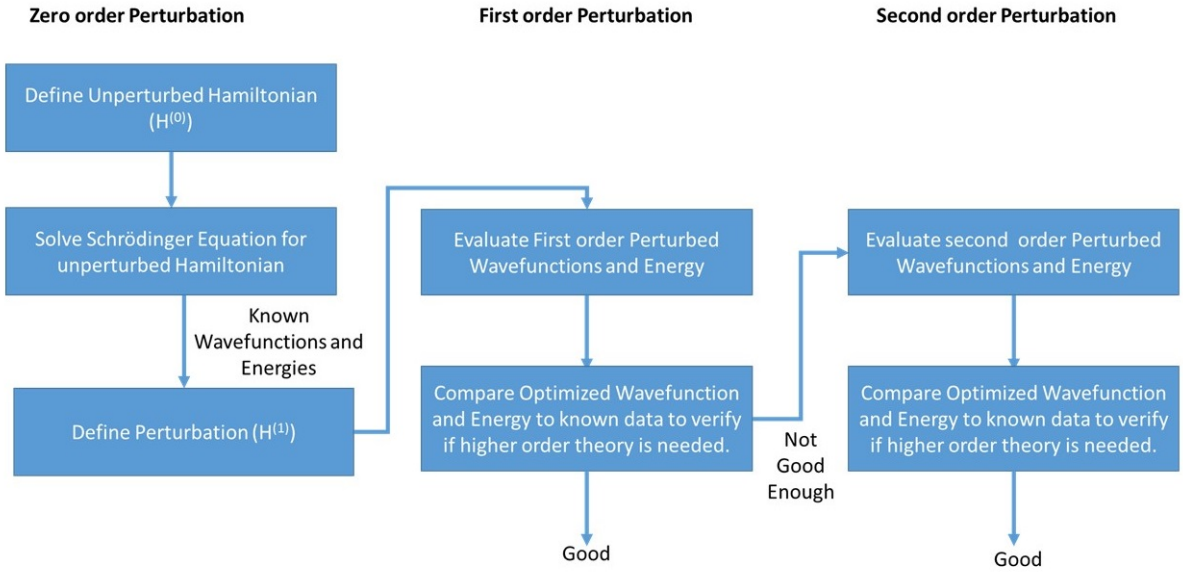


Figure 7.4.1 : Simplified algorithmic flowchart of the Perturbation Theory approximation showing the first two perturbation orders. The process can be continued to third and higher orders. (CC BY 4.0; Delmar Larsen)

Perturbation Theory Does not always Work

It should be noted that there *are* problems that cannot be solved using perturbation theory, even when the perturbation is very weak, although such problems are the exception rather than the rule. One such case is the one-dimensional problem of free particles perturbed by a localized potential of strength λ. Switching on an arbitrarily weak attractive potential causes the $k = 0$ free particle wavefunction to drop below the continuum of plane wave energies and become a localized bound state with binding energy of order λ^2. However, changing the sign of λ to give a repulsive potential there is no bound state, the lowest energy plane wave state stays at energy zero. Therefore the energy shift on switching on the perturbation **cannot** be represented as a power series in λ, the strength of the perturbation.

7.4: Perturbation Theory Expresses the Solutions in Terms of Solved Problems is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **7.4: Perturbation Theory Expresses the Solutions in Terms of Solved Problems** by Delmar Larsen is licensed CC BY-NC-SA 4.0.

7.E: Approximation Methods (Exercises)

Solutions to select questions can be found online.

7.3

Calculate the ground state energy of Harmonic Oscillator using variation method with the following trial wavefunction

$$\phi(x) = |\phi(x)\rangle = \frac{1}{(1+\beta x^2)^2}$$

You may require these definite integrals:

$$\int_{-\infty}^{\infty} \frac{dx}{(1+\beta x^2)^n} = \frac{(2n-3)(2n-5)(2n-7)\ldots(1)}{(2n-2)(2n-4)(2n-6)\ldots(2)} \cdot \pi/\beta^{1/2}$$

$$\int_{-\infty}^{\infty} \frac{dx}{(1-\beta x^2)^n} = \frac{(2n-5)(2n-7)\ldots(1)}{(2n-2)(2n-4)\ldots(2)} \cdot \pi/\beta^{3/2}$$

7.8

What is the variational (trial) energy of the trial function

$$|\phi\rangle = e^{-ax^2}$$

for the ground-state of a harmonic oscillator? Just set up the integral, but do not evaluate. Use

$$\hat{H} = \frac{\hbar^2}{2m}\nabla n^2 + \frac{kx^2}{2}$$

7.9

Use the trial function

$$|\exp^{\frac{-ax^2}{2}}\rangle$$

to set up the integrals to find the ground state energy of a anharmonic oscillator whose potential is $V(x) = cx^5$, but do not evaluate.

7.12

Consider a particle of mass m in a box from $x = -a$ to $x = a$ with $V(x) = -V_0$ for $|x| \geq a$. Assume a trial function of the form

$$|\phi(x)\rangle = l^2 - x^2$$

for $-l < x < l$ and $\psi(x) = 0$ otherwise. l is the parameter. Does the trial function satisfy the requirements of a particle in a box wavefunction?

The result of the variational method was

$$E_\phi(s) = \frac{5}{16}\frac{\hbar^2}{ma^2}\left[\frac{4}{s^2} + \frac{4}{5}\left(8 - \frac{15}{s} + \frac{10}{s^3} - \frac{3}{s^5}\right)\right]$$

Where $s = \frac{l}{a}$ is a new variational parameter for convenience of expression. Derive a polynomial expression for s that can be solved to obtain the value of s that yields the ground state energy, but do not attempt to solve for this value of s.

7.13

Given a trial wavefunction equal to $\sin\lambda(x)$, explain in words a stepwise procedure on how you would go about solving for the energy of this trial wavefunction as well as how to minimize the error.

7.16

Using the variational method approximation, find the ground state energy of a particle in a box using this trial function:

$$|\phi\rangle = N\cos\left(\frac{\pi x}{L}\right)$$

How does is it compare to the true ground state energy?

7.17

For the three-electron detrimental wavefunction

$$\psi = \begin{vmatrix} \phi_A(1) & \phi_A(2) & \phi_A(3) \\ \phi_B(1) & \phi_B(2) & \phi_B(3) \\ \phi_C(1) & \phi_C(2) & \phi_C(3) \end{vmatrix}.$$

confirm that:

a. the interchange of two columns changes the sign of the wavefunction,
b. the interchange of two rows changes the sign of the wavefunction, and
c. the three electrons cannot have the same spin orbital.

7.20

a. What is $\hat{H}^{(0)}$, $\hat{H}^{(1)}$, $\Psi^{(0)}$, and $E^{(0)}$ for an oscillator that has a potential of

$$V(x) = (1/2)kx^2 + x^3 + x^4 + x^5?$$

b. What is $\hat{H}^{(0)}$, $\hat{H}^{(1)}$, $\Psi^{(0)}$, and $E^{(0)}$ for a particle in a box that has a potential of $V(x) = 0$ between 0<x<L?
c. What is $\hat{H}^{(0)}$, $\hat{H}^{(1)}$, $\Psi^{(0)}$, and $E^{(0)}$ for a hydrogenlike atom that has a potential of

$$V(x) = \frac{-e^2}{4\pi\epsilon_o r} + \frac{1}{2}\epsilon r\cos\theta?$$

7.21

Using a harmonic oscillator as the unperturbed problem, calculate the first-order correction to the energy of the $v = 0$ level for the system described as

$$V(x) = \frac{k}{2}x^2 + \frac{m}{6}x^3 + \frac{b}{24}x^4$$

7.22

Using the first order perturbation theory for particle in a box, calculate the ground-state energy for the system

$$V(x) = ax^3 \qquad 0 < x < b$$

https://chem.libretexts.org/@go/page/198701

7.23

In your chemistry lab you were able to manipulate an external electric field to have the strength κ. You're supervisor wants you to figure out what the first-order correction to the ground state energy of a hydrogen like atom of charge N in this electric field.

7.25A

Use first-order perturbation theory to calculate ground-state energy of a harmonic oscillator with a cx^7 added to the end of the potential.

7.25B

In order to calculate the first-order correction to the ground-state energy of the quartic oscillator, use first-order perturbation theory. The potential energy is $V(x) = cx^4$. For this potential use the harmonic oscillator as the unperturbed system. Solve for the perturbing potential as well.

7.26

Solve the following integrals using this trial wavefunction

$$|\phi\rangle = c_1 x(a-x) + c_2 x^2 (a-x)^2$$

For simplicity purposes, we can assume that a = 1.

$$H_{11} = \frac{\hbar^2}{6m} \; S = \frac{1}{30}$$

$$H_{12} = H_{22} = \frac{\hbar^2}{30m} \; S_{12} = S_{21} = \frac{1}{140}$$

$$H_{22} = \frac{\hbar^2}{105m} \; S_{22} = \frac{1}{630}$$

7.27

Use Perturbation Theory to add cubic and quartic perturbations to the SHO and find the first three SHO energy levels. Do this by expanding the Morse potential:

$$V(x) = D(1 - e^{-Bx})^2$$

into polynomials (i.e., a Taylor expansion). Show that the Hamiltonian can be written as

$$\frac{-h^2\nabla^2}{8\pi^2 m} + ax^2 + bx^3 + cx^4$$

Note which terms can be associated with H^0 and which are the H^1 perturbation. What are the relationships between a, b, c, and D, B? How do the new energy levels compare to the old ones?

7.27

Use the perturbation theory to calculate the first - order corrections to the ground state energy of

a. A harmonic oscillator that arises from a cubic and quartic term.
b. A quartic oscillator that arises from only using a quartic term cx^4

and compare the results.

7.E: Approximation Methods (Exercises) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

8: Multielectron Atoms

Electrons with more than one atom, such as Helium (He), and Nitrogen (N), are referred to as multi-electron atoms. Hydrogen is the only atom in the periodic table that has one electron in the orbitals under ground state. We will learn how additional electrons behave and affect a certain atom.

8.1: Atomic and Molecular Calculations are Expressed in Atomic Units
8.2: Perturbation Theory and the Variational Method for Helium
8.3: Hartree-Fock Equations are Solved by the Self-Consistent Field Method
8.4: An Electron Has An Intrinsic Spin Angular Momentum
8.5: Wavefunctions must be Antisymmetric to Interchange of any Two Electrons
8.6: Antisymmetric Wave Functions can be Represented by Slater Determinants
8.7: Hartree-Fock Calculations Give Good Agreement with Experimental Data
8.8: Term Symbols Gives Detailed Descriptions of an Electron Configuration
8.8B: Multi-electron Considerations - A Closer Look at Helium
8.9: The Allowed Values of J - the Total Angular Momentum Quantum Number
8.10: Hund's Rules Determine the Term Symbols of the Ground Electronic States
8.11: Using Atomic Term Symbols to Describe Atomic Spectra
8.E: Multielectron Atoms (Exercises)

Thumbnail: Shell model of Carbon atom. Image used with permission (all right reserved, xponentialdesign).

8: Multielectron Atoms is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

8.1: Atomic and Molecular Calculations are Expressed in Atomic Units

Learning Objectives

- Demonstrate how solving electron structure problems are less cluttered by switching to atomic units instead of SI units.

Atomic units (au or a.u.) form a system of natural units which is especially convenient for atomic physics calculations. Atomic units, like SI units, have a unit of mass, a unit of length, and so on. However, the use and notation is somewhat different from SI. Suppose a particle with a mass of m has 3.4 times the mass of electron. The value of mass m can be written in three ways:

- $m = 3.4\ m_e$: This is the clearest notation (but least common), where the atomic unit is included explicitly as a symbol.
- $m = 3.4\ a.u.$: This notation is ambiguous, but is common. Here, it means that the mass m is 3.4 times the atomic unit of mass. If considering a length L of 3.4 times the atomic unit of length, the equation would look the same, $L = 3.4\ a.u.$ The dimension needs to be inferred from context, which is sloppy.
- $m = 3.4$: This notation is similar to the previous one, and has the same dimensional ambiguity. It comes from formally setting the atomic units to 1 (Table 8.1.1).

This article deals with "Hartree type" of atomic units, where the numerical values of the following four fundamental physical constants are all unity by definition:

Table 8.1.1 : Fundamental atomic units

Dimension	Name	Symbol/Definition	Value in SI units	Value in Atomic Units
mass	electron rest mass	m_e	9.109×10^{-31} kg	1
charge	elementary charge	e	1.602×10^{-19} C	1
action	reduced Planck's constant	$\hbar = \frac{h}{2\pi}$	1.054×10^{-34} J·s	1
electric constant^{-1}	Coulomb force constant	$k_e = \frac{1}{4\pi\epsilon_o}$	8.987 x 10^9 kg·m^3·s^{-2}·C^{-2}	1

Example 8.1.1 : Simplifying the Hamiltonian

Use the atomic units definitions in Table 8.1.1 to contrast the Hamiltonian for a Helium atom in Si units and in atomic units.

Solution

In SI units, the Hamiltonian for a Helium atom is

$$\hat{H} = -\frac{\hbar^2}{2m_e}(\nabla_1^2 + \nabla_2^2) - \frac{2e^2}{4\pi\epsilon_0 r_1} - \frac{2e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}}$$

In atomic units, the same Hamiltonian

$$\hat{H} = -\frac{1}{2}(\nabla_1^2 + \nabla_2^2) - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}$$

All the units that make the SI version of the Hamiltonian disappear to emphasize the key aspects of the operator.

Atomic units are derived from certain fundamental properties of the physical world, and are free of anthropocentric considerations. It should be kept in mind that atomic units were designed for atomic-scale calculations in the present-day universe, with units normalize the reduced Planck constant and also mass and charge of the electron are set to 1, and, as a result, the speed of light in atomic units is a large value, $1/\alpha \approx 137$. For example, the orbital velocity of an electron around a small atom is of the order of 1 in atomic units. Table 8.1.2 give a few derived units. Some of them have proper names and symbols assigned, as indicated in the table.

Table 8.1.2 : Derived atomic units

Dimension	Name	Symbol	Expression	Value in SI units	Value in more common units
length	bohr	a_o	$4\pi\epsilon_0\hbar^2/(m_e e^2) = \hbar/(m_e c\alpha)$	5.291×10^{-11} m	0.052 nm = 0.529 Å
energy	hartree	E_h	$m_e e^4/(4\pi\epsilon_0\hbar)^2 = \alpha^2 m_e c^2$	4.359×10^{-18} J	27.2 eV = 627.5 kcal·mol^{-1}
time			$\hbar/E_h$	2.418×10^{-17} s	
velocity			$a_0 E_h/\hbar = \alpha c$	2.187×10^6 m·s^{-1}	

Bohr model in atomic units

Atomic units are chosen to reflect the properties of electrons in atoms. This is particularly clear from the classical Bohr model of the hydrogen atom in its ground state. The ground state electron orbiting the hydrogen nucleus has (in the classical Bohr model):

- Orbital velocity = 1
- Orbital radius = 1
- Angular momentum = 1
- Orbital period = 2π
- Ionization energy = ½
- Electric field (due to nucleus) = 1
- Electrical attractive force (due to nucleus) = 1

8.1: Atomic and Molecular Calculations are Expressed in Atomic Units is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.1: Atomic and Molecular Calculations are Expressed in Atomic Units** by Wikipedia is licensed CC BY-SA 4.0.

8.2: Perturbation Theory and the Variational Method for Helium

Learning Objectives

- Demonstrate that both perturbation theory and variational methods can be used to solve the electron structure of the helium atom.

Both perturbation theory and variation method (especially the linear variational method) provide good results in approximating the energy and wavefunctions of multi-electron atoms. Below we address both approximations with respect to the helium atom.

Perturbation Theory of the Helium Atom

We use perturbation theory to approach the analytically unsolvable helium atom Schrödinger equation by focusing on the Coulomb repulsion term that makes it different from the simplified Schrödinger equation that we have just solved analytically. The electron-electron repulsion term is conceptualized as a correction, or perturbation, to the Hamiltonian that can be solved exactly, which is called a **zero-order Hamiltonian**. The perturbation term corrects the previous Hamiltonian to make it fit the new problem. In this way the Hamiltonian is built as a sum of terms, and each term is given a name. For example, we call the simplified or starting Hamiltonian, $\hat{H}^0$, the zero order term, and the correction term $\hat{H}^1$.

$$\hat{H} = \hat{H}^0 + \hat{H}^1 \tag{8.2.1}$$

The Hamilonian for the helium atom (in atomic units) is:

$$\hat{H}^0 = \underbrace{-\frac{1}{2}\nabla_1^2 - \frac{2}{r_1}}_{\text{H atom Hamiltonian}} - \underbrace{\frac{1}{2}\nabla_2^2 - \frac{2}{r_2}}_{\text{H atom Hamiltonian}} \tag{8.2.2}$$

$$\hat{H}^1 = \frac{1}{r_{12}} = \frac{1}{|r_1 - r_2|} \tag{8.2.3}$$

The expression for the first-order correction to the energy is

$$\begin{aligned} E^1 &= \langle \psi^0 | \hat{H}^1 | \psi^0 \rangle \\ &= \int \psi^{0*} \hat{H}^1 \psi^0 \, d\tau \end{aligned}$$

Equation ??? is a general expression for the first-order perturbation energy, which provides an improvement or correction to the zero-order energy we already obtained. Hence, E^1 is the average interaction energy of the two electrons calculated using wavefunctions that assume there is no interaction.

The solution to $\hat{H}^0$ (Equation 8.2.2) is the product of two single-electron hydrogen wavefunctions (scaled by the increased nuclear charge) since $\hat{H}^0$ can be separated into independent functions of each electron (i.e., Separation of Variables).

$$|\psi^0\rangle = |\varphi_{1s}(r_1)\varphi_{1s}(r_2)\rangle$$

So the integral in Equation ??? is

$$E^1 = \iint \varphi_{1s}(r_1)\varphi_{1s}(r_2)\frac{1}{r_{12}}\varphi_{1s}(r_1)\varphi_{1s}(r_2)\, d\tau_1 d\tau_2 \tag{8.2.4}$$

where the double integration symbol represents integration over all the spherical polar coordinates of both electrons $r_1, \theta_1, \varphi_1, r_2, \theta_2, \varphi_2$. The evaluation of these six integrals is lengthy. When the integrals are done, the result is E^1 = +34.0 eV so that the total energy calculated using our second approximation method, first-order perturbation theory, is

$$E_{approx2} = E^0 + E^1 = -74.8eV \tag{8.2.5}$$

The new approximate value for the binding energy represents a substantial (~30%) improvement over the zero-order energy:

$$E^0 = \frac{2}{n^2} + \frac{2}{n^2} = 4 \underbrace{E_h}_{hartrees} = 108.8\, eV$$

so the interaction of the two electrons is an important part of the total energy of the helium atom. We can continue with perturbation theory and find the additional corrections, E^2, E^3, etc. For example,

$$E^0 + E^1 + E^2 = -79.2\,eV.$$

So with two corrections to the energy, the calculated result is within 0.3% of the experimental value of -79.01 eV. It takes thirteenth-order perturbation theory (adding E^1 through E^{13} to E^0) to compute an energy for helium that agrees with experiment to within the experimental uncertainty. Interestingly, while we have improved the calculated energy so that it is much closer to the experimental value, we learn nothing new about the helium atom wavefunction by applying the first-order perturbation theory to the energy above. He need to expand the wavefunctions to first order perturbation theory, which requires more effort. Below, we will employ the variational method approximation to modify zero-order wavefunctions to address one of the ways that electrons are expected to interact with each other.

The Hartree Unit of Energy

The hartree is the atomic unit of energy (named after the British physicist Douglas Hartree) and is defined as

$$E_h = 2R_H hc$$

where R_H is the Rydberg constant, h is the Planck constant and c is the speed of light.

$$\begin{aligned} E_h &= 4.359 \times 10^{-18}\,J \\ &= 27.21\,eV. \end{aligned}$$

The hartree is usually used as a unit of energy in atomic physics and computational chemistry. As discussed before for hydrogen emission, IR, and microwave spectroscopies, experimental measurements prefer the electronvolt (eV) or the wavenumber (cm^{-1}).

Variational Method Applied to the Helium Method

As discussed in Section 6.7, because of the electron-electron interactions, the Schrödinger's Equation cannot be solved exactly for the helium atom or more complicated atomic or ionic species. However, the ground-state energy of the helium atom can be estimated using approximate methods. One of these is the variational method which requires the minimizing of the following variational integral.

$$E_{trial} = \frac{\langle \psi_{trial} | \hat{H} | \psi_{trial} \rangle}{\langle \psi_{trial} | \psi_{trial} \rangle} \tag{8.2.6}$$

$$= \frac{\int_0^\infty \psi_{trial}^* \hat{H} \psi_{trial} d\tau}{\int_0^\infty \psi_{trial}^2 \, d\tau} \tag{8.2.7}$$

The five trial wavefunctions discussions below are equally "valid" trial wavefunctions that describe the probability of finding each electron (technically the wavefunction squared). What separates the "poor" approximations from the "good" approximation is whether the trial wavefunction predicts experimental results. Consequently, for all the approximations used for the rest of this TextMap, it is important to compare the theoretical results to the "true" (i.e., experimental) results. No matter how complicated an approximation is, it is only as good as the accuracy of its predicted values to experimental values.

Trial Wavefunction #1: Simple Orbital Approximation with One Parameter

As is clear from Equation 8.2.6, the variational method approximation requires that a trial wavefunction with one or more adjustable parameters be chosen. A logical first choice for such a multi-electron wavefunction would be to assume that the electrons in the helium atom occupy two identical, but scaled, hydrogen 1s orbitals.

$$|\psi(1,2)\rangle_{trial} = \phi(1)\phi(2) \tag{8.2.8}$$

$$= \exp\left[-\alpha(r_1 + r_2)\right] \tag{8.2.9}$$

The variational energy obtained after minimizing Equation 8.2.7 after substituting the trial wavefunction (Equation 8.2.9) by varying α is

$$E_{trial} = -2.84766\ E_h$$

and the experimentally determined ground-state energy for the helium atom is the sum of first and second ionization energies

$$E_{\text{exp}} = I_1 + I_2 = -2.90372\ E_h \tag{8.2.10}$$

The deviation of energy for the optimized trial wavefunction from the experimental value is

$$\left|\frac{E_{trial}(\alpha) - E_{\text{exp}}}{E_{\text{exp}}}\right| = \left|\frac{-2.84766\ E_h + 2.90372\ E_h}{-2.90372\ E_h}\right| \tag{8.2.11}$$

$$= 1.93\% \tag{8.2.12}$$

The value of -2.8477 hartrees is within 2% of the known ground-state energy of the helium atom. The error in the calculation is attributed to the fact that the wavefunction is based on the orbital approximation and, therefore, does not adequately take electron-electron interactions into account. In other words, this wavefunction gives the electrons too much independence, given that they have like charges and tend to avoid one another.

Trial Wavefunction #2: Orbital Approximation with Two Parameters

Some electron-electron interactions can be built into the multi-electron wavefunction by assuming that each electron is in an orbital which is a linear combination of two different and scaled hydrogen 1s orbitals.

$$\phi(r_1) = \exp(-\alpha r_1) + \exp(-\beta r_1) \tag{8.2.13}$$

Under the orbital approximation this assumption gives a trial wavefunction of the form

$$|\psi(1,2)\rangle_{trial} = \phi(1)\phi(2) \tag{8.2.14}$$

$$= \exp(-\alpha r_1)\exp(-\alpha r_2) + \exp(-\alpha r_1)\exp(-\beta r_2) + \exp(-\beta r_1)\exp(-\alpha r_2) + \exp(-\beta r_1)\exp(-\beta r_2) \tag{8.2.15}$$

Inspection of this trial wavefunction indicates that 50% of the time the electrons are in different orbitals, while for the first trial wavefunction the electrons were in the same orbital 100% of the time. Notice the enormous increase in the complexity of the variational expression for the energy for this trial wavefunction (Equation 8.2.7). However, the calculation is very similar to that using the previous trial wavefunction. The differences are that in this case the expression for the energy is more complex and that it is being minimized simultaneously with respect to two parameters (α and β) rather than just one (α).

The variational energy obtained after minimizing Equation 8.2.7 after substituting the trial wavefunction (Equation 8.2.15) by varying α and β is

$$E_{trial} = -2.86035\ E_h$$

The deviation of energy for the optimized trial wavefunction from the experimental value (Equation 8.2.10) is

$$\left|\frac{E_{trial}(\alpha,\beta) - E_{\text{exp}}}{E_{\text{exp}}}\right| = \left|\frac{-2.86035\ E_h + 2.90372\ E_h}{-2.90372\ E_h}\right| \tag{8.2.16}$$

$$= 1.49\% \tag{8.2.17}$$

Clearly introducing some electron-electron interactions into the trial wavefunction has improved the agreement between theory and experiment (Equation 8.2.12 vs. 8.2.17).

Trial Wavefunction #3: Orbital Approximation with Two Parameters

The extent of electron-electron interactions can be increased further by eliminating the first and last term in the second trial wavefunction (Equation 8.2.15). This yields a multi-electron wavefunction of the form,

$$|\psi(1,2)\rangle_{trial} = \exp(-\alpha r_1)\exp(-\beta r_2) + \exp(-\beta r_1)\exp(-\alpha r_2) \tag{8.2.18}$$

This trial wavefunction places the electrons in different scaled hydrogen 1s orbitals 100% of the time this adds further improvement in the agreement with the literature value of the ground-state energy is obtained. The variational energy obtained after

minimizing Equation 8.2.7 after substituting the trial wavefunction (Equation 8.2.18) by varying α and β is

$$E_{trial} = -2.87566\ E_h$$

The deviation of energy for the optimized trial wavefunction from the experimental value (Equation 8.2.10) is

$$\left|\frac{E_{trial}(\alpha,\beta) - E_{\text{exp}}}{E_{\text{exp}}}\right| = \left|\frac{-2.87566\ E_h + 2.90372\ E_h}{-2.90372\ E_h}\right| \tag{8.2.19}$$

$$= 0.97\% \tag{8.2.20}$$

This result is within 1% of the actual ground-state energy of the helium atom.

Trial Wavefunction #4: Approximation with Two Parameters

The third trial wavefunction, however, still rests on the orbital approximation and, therefore, does not treat electron-electron interactions adequately. Hylleraas took the calculation a step further by introducing electron-electron interactions directly into the first trial wavefunction by adding a term, r_{12}, involving the inter-electron separation.

$$|\psi_{trial}(1,2)\rangle = (\exp[-\alpha(r_1 + r_2)])\,(1 + \beta r_{12}) \tag{8.2.21}$$

In the trial multi-electron wavefunction of Equation 8.2.21, if the electrons are far apart, then r_{12} is large and the magnitude of the wavefunction increases to favor that configuration. The variational energy obtained after minimizing Equation 8.2.7 after substituting the trial wavefunction (Equation 8.2.21) by varying α and β is

$$E_{trial} = -2.89112\ E_h$$

The deviation of energy for the optimized trial wavefunction from the experimental value (Equation 8.2.10) is

$$\left|\frac{E_{trial}(\alpha,\beta) - E_{\text{exp}}}{E_{\text{exp}}}\right| = \left|\frac{-2.89112\ E_h + 2.90372\ E_h}{-2.90372\ E_h}\right| \tag{8.2.22}$$

$$= 0.43\% \tag{8.2.23}$$

This modification of the trial wavefunction has further improved the agreement between theory and experiment to within 0.5%.

Fifth Trial Wavefunction #5: Approximation with Three Parameters

Chandrasakar brought about further improvement by adding Hylleraas's r_{12} term to the third trial wavefunction (Equation 8.2.18) as shown here.

$$|\psi(1,2)\rangle_{trial} = [\exp(-\alpha r_1)\exp(-\beta r_2) + \exp(-\beta r_1)\exp(-\alpha r_2)]\,[1 + \gamma r_{12}] \tag{8.2.24}$$

Chandrasakar's three parameter wavefunction gives rise to a fairly complicated variational expression for ground-state energy. The variational energy obtained after minimizing Equation 8.2.7 after substituting the trial wavefunction (Equation 8.2.24) by varying α, β and γ is

$$E_{trial} = -2.90143\ E_h$$

The deviation of energy for the optimized trial wavefunction from the experimental value (Equation 8.2.10) is

$$\left|\frac{E_{trial}(\alpha,\beta,\gamma) - E_{\text{exp}}}{E_{\text{exp}}}\right| = \left|\frac{-2.90143 + 2.90372\ E_h}{-2.90372\ E_h}\right| \tag{8.2.25}$$

$$= 0.0789\% \tag{8.2.26}$$

Chandrasakar's wavefunction gives a result for helium that is within 0.07% of the experimental value for the ground-state energy.

Summary

The purpose of this section is to examine five trial wavefunctions for the helium atom used within the Perturbation Theory and Variational method approximation. For the Variational method approximation, the calculations begin with an uncorrelated wavefunction in which both electrons are placed in a hydrogenic orbital with scale factor α. The next four trial functions use several methods to increase the amount of electron-electron interactions in the wavefunction. As the summary of results that is appended shows this gives increasingly more favorable agreement with the experimentally determined value for the ground-state

energy of the species under study. The detailed calculations show that the reason for this improved agreement with experiment is due to a reduction in electron-electron repulsion.

Five variational method calculations that have been outlined above for the helium atom ($Z = 2$) can be repeated for two-electron atoms (e.g., H^-, Li^+, Be^{2+}, etc). The hydride anion is a particularly interesting case because the first two trial wavefunctions do not predict a stable ion (i.e., they are poor approximations). This indicates that electron-electron interactions is an especially important issue for atoms and ions with small nuclear charge.

8.2: Perturbation Theory and the Variational Method for Helium is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.2: Perturbation Theory and the Variational Method for Helium** by Frank Rioux is licensed CC BY 4.0. Original source: https://faculty.csbsju.edu/frioux/workinprogress.html.

8.3: Hartree-Fock Equations are Solved by the Self-Consistent Field Method

Learning Objectives

- Show how the Hartree approximation can be used to solve for the wavefunctions and energies of multi-electron atoms.
- Understand the orbital approach of independent orbitals is an approximation to a mutlti-electron system with the motions of all electrons coupled together.
- Demonstate how the Self-Consistant Field (SCF) calculation is needed to support the Hartree approximation

The Hartree method is used to approximate the wavefunction and the energy of a quantum multi-electron system in a stationary state. This approximation assumes that the exact N-body wavefunction of the system can be approximated by a product of single-electron wavefunctions. By invoking the variational method, one can derive a set of N-coupled equations for the N spin orbitals. A solution of these equations yields the Hartree wavefunction and energy of the system. It is one step better than the "Ignorance is Bliss" approach, discussed previously, but still far from modern state-off-the-art methods.

Unsolvable Systems

The Hamiltonian for a generic multielectron atom includes nucleus-electron attraction terms for the additional electrons with a general charge Z; e.g.

$$V_{\text{nuclear-electron}}(r_1) = -\frac{Z}{|\mathbf{r}-\mathbf{R}|} \tag{8.3.1}$$

in atomic units with $|\mathbf{r}-\mathbf{R}|$ is the distance between the electron and the nucleus, The Hamiltonian must also have terms for electron-electron repulsion (also in atomic units)

$$V_{\text{electron-electron}}(r_{12}) = \frac{1}{|\mathbf{r}-\mathbf{r}'|}$$

with $|\mathbf{r}-\mathbf{r}'|$ is the distance between electron 1 and electron 2. So the proper multi-electron Hamiltonian can be constructed

$$\hat{H}(r_1, r_2, \ldots r_n) = -\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 + \sum_i V_{\text{nuclear-electron}}(r_i) + \sum_{i\neq j} V_{\text{electron-electron}}(r_{ij}) \tag{8.3.2}$$

Given what we have learned from the previous quantum mechanical systems we've studied, we predict that exact solutions to the multi-electron Schrödinger equation would consist of a family of multi-electron wavefunctions, each with an associated energy eigenvalue. These wavefunctions and energies would describe the ground and excited states of the multi-electron atom, just as the hydrogen wavefunctions and their associated energies describe the ground and excited states of the hydrogen atom. We would predict quantum numbers to be involved, as well.

The fact that electrons interact through their electron-electron repulsion (final term in Equation 8.3.2) means that an exact wavefunction for a multi-electron system would be a single function that depends **simultaneously** upon the coordinates of all the electrons; i.e., a multi-electron wavefunction:

$$|\psi(r_1, r_2, \cdots r_i)\rangle \tag{8.3.3}$$

Unfortunately, the electron-electron repulsion terms make it impossible to find an exact solution to the Schrödinger equation for many-electron atoms.

The Hartree Approximation

The method for finding best possible one-electron wavefunctions that was published by Douglas Hartree in 1948 and improved two years later by Vladimir Fock. For the Schrödinger equation to be analytically solvable, the variables must be separable - the variables are the coordinates of the electrons. To separate the variables in a way that retains information about electron-electron interactions, the electron-electron term (Equation 8.3.1) must be approximated so it depends only on the coordinates of one electron. Such an approximate Hamiltonian can account for the interaction of the electrons in **an average way**. The exact one-

electron eigenfunctions of this approximate Hamiltonian then can be found by solving the Schrödinger equation. These functions are the best possible **one-electron functions**.

The Hartree approximation starts by invoking an initial *ansatz* that the multi-electron wavefunction in Equation 8.3.3 can be expanded as a product of single-electron wavefunctions (i.e., orbitals)

$$|\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)\rangle \approx \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\ldots\psi_N(\mathbf{r}_N)$$

from which it follows that the electrons are *independent*, and interact only via the mean-field Coulomb potential. This yields one-electron Schrödinger equations of the form

$$-\frac{\hbar^2}{2m}\nabla^2\psi_i(\mathbf{r}) + V(\mathbf{r})\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r})$$

or

$$H_e(r)\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r})$$

where $V(r)$ is the potential in which the electron moves; this includes both the nuclear-electron interaction

$$V_{nucleus}(\mathbf{r}) = -Ze^2\sum_R \frac{1}{|\mathbf{r}-\mathbf{R}|}$$

and the mean field arising from the $N-1$ other electrons. We smear the other electrons out into a smooth negative charge density $\rho(\mathbf{r}')$ leading to a potential of the form

$$V_{electron}(\mathbf{r}) = -e\int d\mathbf{r}'\rho(\mathbf{r}')\frac{1}{|\mathbf{r}-\mathbf{r}'|}$$

where

$$\rho(\mathbf{r}) = \sum_i^{\text{occupied}} |\psi(\mathbf{r})|^2.$$

The sum over runs over all occupied states; i.e., only the states of electrons that exist in the atom. The wavefunctions that from this approach with the Hamiltonian $H_e(r)$ involve possess three kinds of energies discussed below.

Three Energies within the Hartree Approximation

The total energy ϵ_j of the orbital ϕ_j, is the sum of the above three contributions:

1. **Kinetic Energy:** The Kinetic energy of the electron has an average value is computed by taking the expectation value of the kinetic energy operator

   $$\frac{-\hbar^2}{2m}\nabla^2$$

   with respect to any particular solution $\phi_j(r)$ to the Schrödinger equation:

   $$KE = \langle\phi_j|\frac{-\hbar^2}{2m}\nabla^2|\phi_j\rangle$$

2. **Nuclear-Electron Coulombic Attraction Energy:** Coulombic attraction energy with the nucleus of charge Z:

   $$\langle\phi_j|\frac{-Ze^2}{|\mathbf{r}-\mathbf{R}|}|\phi_j\rangle$$

3. **Electron-Electron Coulombic Repulsion Energy:** Coulomb repulsion energies with all of the $N-1$ other electrons, which are assumed to occupy other atomic orbitals denoted ϕ_K, with this energy computed as

   $$\sum_{j\neq k}\langle\phi_j(r)\phi_k(r')|\frac{e^2}{|r-r'|}|\phi_j(r)\phi_k(r')\rangle. \tag{8.3.4}$$

The Dirac notation $\langle\phi_j(r)\phi_k(r')|\frac{e^2}{|r-r'|}|\phi_j(r)\phi_k(r')\rangle$ is used to represent the two-electron (six-dimensional) Coulomb integral

$$J_{j,k} = \int |\phi_j(r)|^2 |\phi_k(r')|^2 \frac{e^2}{r-r'} dr dr'$$

that describes the Coulomb repulsion between the charge density $|\phi_j(r)|^2$ for the electron in ϕ_j and the charge density $|\phi_k(r')|^2$ for the electron in ϕ_k. Of course, the sum over k must be limited to exclude $k = j$ to avoid counting a "self-interaction" of the electron in orbital ϕ_j with itself.

Adding these all together to get the total energy ϵ_j of the orbital ϕ_j:

$$\epsilon_J = \langle\phi_j|\frac{-\hbar^2}{2m}\nabla^2|\phi_j\rangle + \langle\phi_j|\frac{-Ze^2}{|\mathbf{r}-\mathbf{R}|}|\phi_j\rangle + \sum_{j\neq k}\langle\phi_j(r)\phi_k(r')|\frac{e^2}{|r-r'|}|\phi_j(r)\phi_k(r')\rangle.$$

This treatment of the electrons and their orbitals is referred to as the **Hartree-level of theory**.

When screened hydrogenic atomic orbitals are used to approximate the ϕ_j and ϕ_K orbitals, the resultant ϵ_J values do **not** produce accurate predictions. For example, the negative of ϵ_J should approximate the ionization energy for removal of an electron from the orbitals ϕ_j. Such ionization potentials (IP s) can be measured, and the measured values do not agree well with the theoretical values when a crude screening approximation is made for the atomic orbitals.

The Self-Consistant Field (SCF) Approach to the Variational Method

The Hartree Equations are nonlinear and must be solved iteratively. This is because if particles interact, that interaction must be in the Hamiltonian. So until we know where the particles are, we cannot write down the Hamiltonian, but until we know the Hamiltonian, we cannot tell where the particles are.

The idea is to solve the Schrödinger equation for an electron moving in the potential of the nucleus and all the other electrons. We start with a guess for the trial electron charge density, solve N/2 one-particle Schrödinger equations (initially identical) to obtain N electron wavefunctions. Then we construct the potential for each wavefunction from that of the nucleus and that of all the other electrons, symmetrize it, and solve the N/2 Schrödinger equations again. This method is ideal for a computer, because it is easily written as an algorithm (Figure 8.3.1).

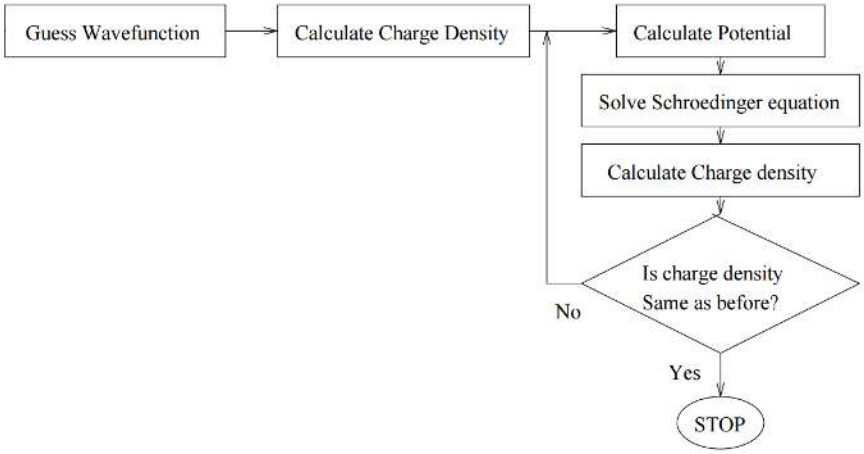


Figure 8.3.1 : Algorithm for Self-consistent field theory.

Although we are concerned here with atoms, the same methodology is used for molecules or even solids (with appropriate potential symmetries and boundary conditions). This is a variational method, so wherever we refer to wavefunctions, we assume that they are expanded in some appropriate basis set.

> *Fock improved on Hartree's method by using proper "antisymmetrized wavefunctions" (called the Hartree-Fock method) instead of simple one-electron wavefunctions.*

Shielding and Effective Charges Concepts are Useful

The hydrogen-like orbitals provide qualitative descriptions of orbitals of atoms with more than a single electron. By introducing the concept of screening as a way to represent the repulsive interactions among the electrons of an atom, an effective nuclear charge Z_{eff} can be used in place of Z in the hydrogenic $\psi_{n,l,m}$ and $E_{n,l}$ formulas to generate approximate atomic orbitals to be filled by electrons in a many-electron atom. For example, in the crudest approximation of a carbon atom, the two $1s$ electrons experience the full nuclear attraction so $Z_{\text{eff}} = 6$ for them, whereas the $2s$ and $2p$ electrons are screened by the two $1s$ electrons, so $Z_{\text{eff}} = 4$ for them. Within this approximation, one then occupies two $1s$ orbitals with $Z = 6$, two $2s$ orbitals with $Z = 4$ and two $2p$ orbitals with $Z = 4$ in forming the full six-electron product wavefunction of the lowest-energy state of carbon

$$|\psi(1,2,\ldots,6)\rangle = |\psi_{1s}(1)\psi_{1s}(2)\psi_{2s}(3)\ldots\psi_{1p}(6)\rangle. \tag{8.3.5}$$

However, such approximate orbitals are not sufficiently accurate to be of use in quantitative simulations of atomic and molecular structure. In particular, their energies do not properly follow the trends in atomic orbital (AO) energies that are taught in introductory chemistry classes (Figure 8.3.2).

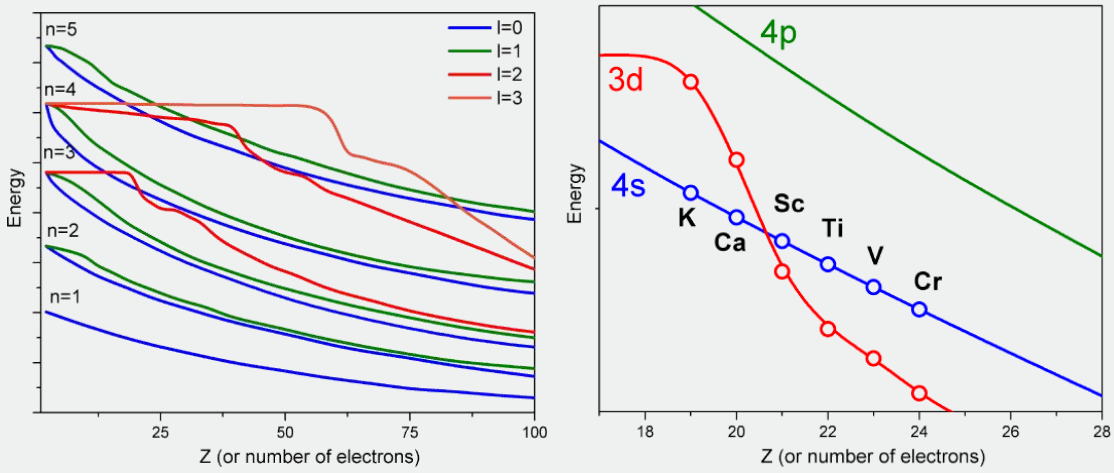


Figure 8.3.2 : Energies of Atomic Orbitals as Functions of the number of electrons in Neutral Atoms. (CC BY 4.0; Delmar Larsen).

For example, the relative energies of the $3d$ and $4s$ orbitals are not adequately described in a model that treats electron repulsion effects in terms of a simple screening factor. So, now it is time to examine how we can move beyond the screening model and take the electron repulsion effects, which cause the inter-electronic couplings that render the Schrödinger equation insolvable, into account in a more reliable manner.

8.3: Hartree-Fock Equations are Solved by the Self-Consistent Field Method is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.3: Hartree-Fock Equations are Solved by the Self-Consistent Field Method** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

8.4: An Electron Has An Intrinsic Spin Angular Momentum

Learning Objectives

- Understand the forth quantum number for electrons - spin.
- Understand how spin is connected to magnetic properties of the electrons and atoms.
- Understand how to break degeneracy via externally applied magnetic fields in electrons and atoms.

Imagine doing a hypothetical experiment that would lead to the discovery of electron spin. Your laboratory has just purchased a microwave spectrometer with variable magnetic field capacity. We try the new instrument with hydrogen atoms using a magnetic field of 10^4 Gauss and look for the absorption of microwave radiation as we scan the frequency of our microwave generator (Figure 8.4.1).

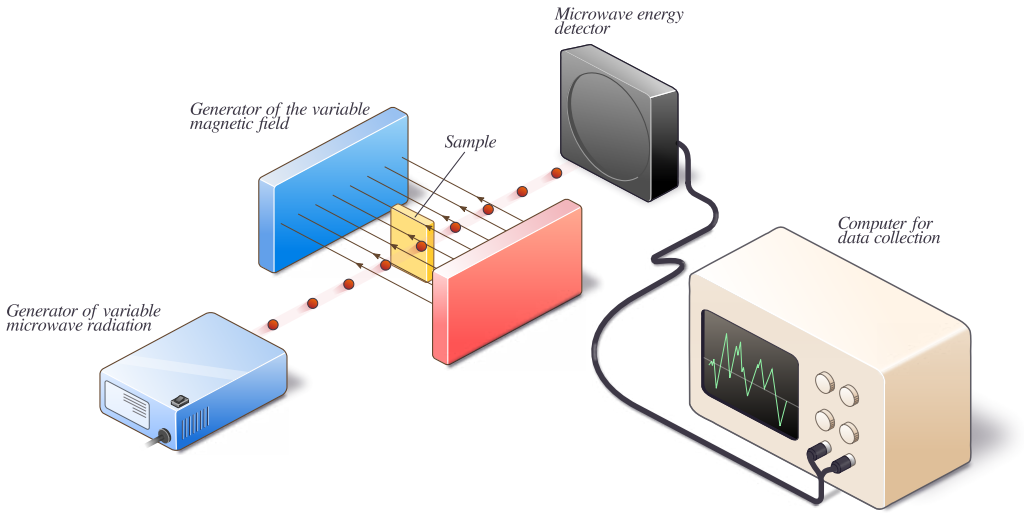


Figure 8.4.1 : Schematic diagram of a microwave spectrometer with the sample in a variable magnetic field. The strength of the magnetic field is set, and the sample's absorption of microwave photons is measured for a range of microwave photon energies (or frequencies). (CC BY-NC 4.0; Ümit Kaya via LibreTexts)

Finally we see absorption at a microwave photon frequency of $28 \times 10^9\ Hz$ (28 gigahertz). This result is really surprising from several perspectives. Each hydrogen atom is in its ground state, with the electron in a 1s orbital. The lowest energy electronic transition that we predict based on existing theory (the electronic transition from the ground state (ψ_{100} to ψ_{21m}) requires an energy that lies in the vacuum ultraviolet (the lower Lyman line at 121 nm), not the microwave, region of the spectrum. Furthermore, when we vary the magnetic field we note that the frequency at which the absorption occurs varies in proportion to the magnetic field.

The Zeeman Effect: Breaking Degeneracies with Magnetic Fields

Magnetism results from the circular motion of charged particles. This property is demonstrated on a macroscopic scale by making an electromagnet from a coil of wire and a battery. Electrons moving through the coil produce a magnetic field, which can be thought of as originating from a magnetic dipole or a bar magnet. Electrons in atoms also are moving charges with angular momentum so they too produce a magnetic dipole, which is why some materials are magnetic. A **magnetic dipole** interacts with a magnetic field, and the energy of this interaction is given by the scalar product of the magnetic dipole moment, and the magnetic field, $\vec{B}$.

$$E_B = -\vec{\mu}_m \cdot \vec{B} \tag{8.4.1}$$

Pieter Zeeman was one of the first to observe the splittings of spectral lines in a magnetic field caused by this interaction. Consequently such splittings are known as the Zeeman effect. (Figure 8.4.2). The expectation value calculated for the total energy in this case is the sum of the energy in the absence of the field, E_n, plus the Zeeman energy:

$$\langle E \rangle = E_n + \frac{e\hbar B_z m_l}{2m_e} \tag{8.4.2}$$

$$= E_n + \mu_B B_z m_l \tag{8.4.3}$$

The factor

$$\frac{e\hbar}{2m_e} = -\gamma_e \hbar = \mu_B \tag{8.4.4}$$

defines the constant μ_B, called the **Bohr magneton**, which is taken to be the fundamental magnetic moment. It has units of 9.2732×10^{-21} erg/Gauss or 9.2732×10^{-24} Joule/Tesla. This factor will help you to relate magnetic fields, measured in Gauss or Tesla, to energies, measured in ergs or Joules, for any particle with a charge and mass the same as an electron.

Equation 8.4.3 demonstrates that that m_l quantum number degeneracy of the hydrogen atom is removed by the externally applied magnetic field. For example, the three hydrogen atom eigenstates $|\psi_{211}\rangle$, $|\psi_{21-1}\rangle$, and $|\psi_{210}\rangle$ are degenerate in zero magnetic field, but have different energies in an externally applied magnetic field (Figure 8.4.2).

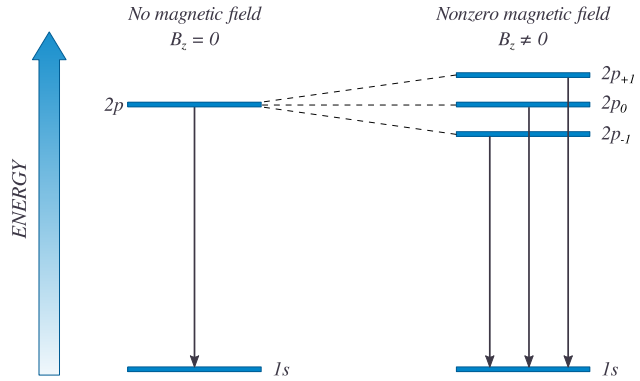


Figure 8.4.2 : The Zeeman effect for the $2s \rightarrow 1s$ transition. Emission when an electron switches from a 2p orbital to a 1s orbital occurs at only one energy in the absence of a magnetic field, but can occur at three different energies in the presence of a magnetic field. (CC BY-NC; Ümit Kaya via LibreTexts)

The $m_l = 0$ state, for which the component of angular momentum and hence also the magnetic moment in the external field direction is zero, experiences no interaction with the magnetic field. The $m_l = +1$ state, for which the angular momentum in the z-direction is +ħ and the magnetic moment is in the opposite direction, against the field, experiences a raising of energy in the presence of a field. Maintaining the magnetic dipole against the external field direction is like holding a small bar magnet with its poles aligned exactly opposite to the poles of a large magnet. It is a higher energy situation than when the magnetic moments are aligned with each other.

Electron Spin and the Stern-Gerlach Experiment

To discover new things, experimentalists sometimes must explore new areas in spite of contrary theoretical predictions. Our theory of the hydrogen atom at this point gives no reason to look for absorption in the microwave region of the spectrum. By doing the crazy experiment outlines above, we discovered that when an electron is in the $|1s\rangle$ orbital of the hydrogen atom, there are two different states that have the same energy. When a magnetic field is applied, this degeneracy is removed, and microwave radiation can cause transitions between the two states. In the rest of this section, we see what can be deduced from this experimental observation. This experiment actually could be done with electron spin resonance spectrometers available today (Figure 8.4.1). To explain our observations, a new model for the hydrogen atom. Our original model for the hydrogen atom accounted for the motion of the electron and proton in our three-dimensional world; the new model needs something else that can give rise to an additional Zeeman-like effect. We need a charged particle with angular momentum to produce a magnetic moment, just like that obtained by the orbital motion of the electron. We can postulate that our observation results from a motion of the electron that was not considered in the last section - electron spin. We have a charged particle spinning on its axis. We then have charge moving in a circle, angular momentum, and a magnetic moment, which interacts with the magnetic field and gives us the Zeeman-like effect that we observed (Figure 8.4.2).

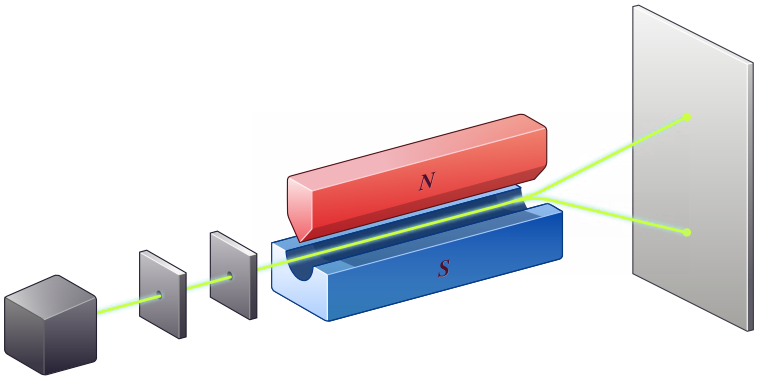


Figure 8.4.2 : Stern–Gerlach experiment: silver atoms travel through an inhomogeneous magnetic field and are deflected up or down depending on their spin. (CC BY-NC 4.0; Ümit Kaya via LibreTexts)

In 1920, Otto Stern and Walter Gerlach designed an experiment, which unintentionally led to the discovery that electrons have their own individual, continuous spin even as they move along their orbital of an atom. Today, this electron spin is indicated by the fourth quantum number, also known as the **Electron Spin Quantum Number** and denoted by m_s. In 1925, Samuel Goudsmit and George Uhlenbeck made the claim that features of the hydrogen spectrum that were unexamined might by explained by assuming electrons act as if it has a spin, which can be denoted by an arrow pointing up, which is +1/2, or an arrow pointing down, which is -1/2. The Stern and Gerlach experiment which demonstrated this was done with a beam of vaporized silver atoms that split into two beams after passing through a magnetic field (Figure 8.4.2). An explanation of this is that an electron has a magnetic field due to its spin. When electrons that have opposite spins are put together, there is no net magnetic field because the positive and negative spins cancel each other out. The silver atom used in the experiment has a total of 47 electrons, 23 of one spin type, and 24 of the opposite. Because electrons of the same spin cancel each other out, the *one unpaired electron* in the atom will determine the spin.

Spin Eigenstates and Eigenvalues

To describe electron spin from a quantum mechanical perspective, we must have spin wavefunctions and spin operators. The properties of the spin states are deduced from experimental observations and by analogy with our treatment of the states arising from the orbital angular momentum of the electron. The important feature of the spinning electron is the spin angular momentum vector, which we label $\boldsymbol{S}$ by analogy with the orbital angular momentum $\boldsymbol{L}$. We define spin angular momentum operators with the same properties that we found for the rotational and orbital angular momentum operators. After all, angular momentum is angular momentum, no matter if it is *orbital* or *spin* in nature.

We found that

$$\hat{L}^2|Y_l^{m_l}\rangle = l(l+1)\hbar^2|Y_l^{m_l}\rangle$$

so by analogy for the spin states, we must have

$$\hat{S}^2|\sigma_s^{m_s}\rangle = s(s+1)\hbar^2|\sigma_s^{m_s}\rangle$$

where σ is a spin wavefunction with quantum numbers s and m_s that obey the same rules as the quantum numbers l and m_l associated with the spherical harmonic wavefunction Y. We also found the project of the orbital angular momentum on the z-axis is

$$\hat{L}_z|Y_l^{m_l}\rangle = m_l\hbar|Y_l^{m_l}\rangle$$

so by analogy, we must have a similar projection for the spin angular momentum:

$$\hat{S}_z|\sigma_s^{m_s}\rangle = m_s\hbar|\sigma_s^{m_s}\rangle \tag{8.4.5}$$

Since m_l ranges in integer steps from $-l$ to $+l$, also by analogy m_s ranges in integer steps from $-s$ to $+s$. In our hypothetical experiment, we observed one absorption transition, which means there are two spin states. Consequently, the two values of m_s must

be $+s$ and $-s$, and the difference in m_s for the two states, labeled f and i below, must be the smallest integer step, i.e., 1. The result of this logic is that

$$\begin{aligned} m_{s,f} - m_{s,i} &= 1 \\ (+s) - (-s) &= 1 \\ 2s &= 1 \\ s &= \frac{1}{2} \end{aligned} \tag{8.4.6}$$

Therefore our conclusion is that the magnitude of the spin quantum number is 1/2 and the values for m_s are +1/2 and -1/2. The two spin states correspond to spinning clockwise and counter-clockwise with positive and negative projections of the spin angular momentum onto the z-axis (Figure 8.4.3). The state with a positive projection, m_s = +1/2, is called α; the other is called β. These spin states are arbitrarily labeled α and β, and the associated spin wavefunctions also are designated by $|\alpha\rangle$ and $|\beta\rangle$.

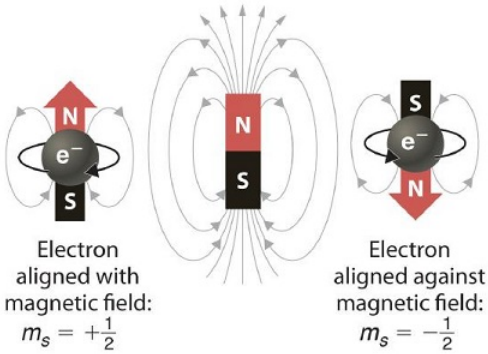


Figure 8.4.3 : Electron Spin. In a magnetic field, an electron has two possible orientations with different energies, one with spin up, aligned with the magnetic field, and one with spin down, aligned against it. All other orientations are forbidden. (CC BY-SA-NC 3.0; Anonymous).

From Equation 8.4.5, the magnitude of the z-component of spin angular momentum, S_z, is given by

$$S_z = m_s \hbar \tag{8.4.7}$$

so the value of S_z is $+\hbar/2$ for spin state α and $-\hbar/2$ for spin state β. Hence, we conclude that the α spin state, where the magnetic moment is aligned against the external field direction, has a **greater energy** than the β spin state.

Electrons are not Actually Spinning

Electron's hypothetical surface would have to be moving *faster* than the speed of light for it to rotate quickly enough to produce the observed angular momentum. Hence, an electron is not simply a spinning ball or ring and electron spin appears to be an *intrinsic angular moment* of the particle rather than a consequence of the rotation of a charge particle like Figure 8.4.3 suggests. Despite this, the term "electron spin" persists in quantum vernacular.

Properties of Spin Eigenstates

Even though we do not know their functional forms, the spin wavefunctions are taken to be normalized and orthogonal to each other.

$$\int \alpha^* \alpha \, d\tau_s = \int \beta^* \beta \, d\tau_s = 1 \tag{8.4.8}$$

or in braket notation

$$\langle \alpha | \alpha \rangle = \langle \beta | \beta \rangle = 1 \tag{8.4.9}$$

and

$$\int \alpha^* \beta \, d\tau_s = \int \beta^* \alpha \, d\tau_s = 0 \tag{8.4.10}$$

or in braket notation

$$\langle\alpha|\beta\rangle = \langle\alpha|\beta\rangle = 0 \tag{8.4.11}$$

where the integral is over the spin variable τ_s.

Now let's apply these deductions to the experimental observations in our hypothetical microwave experiment in Figure 8.4.1 . We can account for the frequency of the transition (ν= 28 gigahertz) that was observed in this hypothetical experiment in terms of the magnetic moment of the spinning electron and the strength of the magnetic field. The photon energy, $h\nu$, is given by the difference between the energies of the two states, E_α and E_β

$$\Delta E = h\nu \tag{8.4.12}$$

$$= E_\alpha - E_\beta \tag{8.4.13}$$

The energies of these two states consist of the sum of the energy of an electron in a 1s orbital, E_{1s}, and the energy due to the interaction of the spin magnetic dipole moment of the electron, μ_s, with the magnetic field, B.

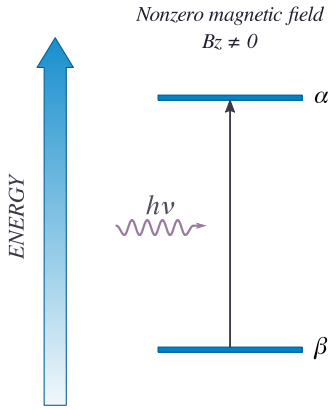


Figure 8.4.4 : Absorption of a photon to cause a transition from the β to the α state. (CC BY-NC; Ümit Kaya via LibreTexts)

The two states with distinct values for spin magnetic moment μ_s are denoted by the subscripts α and β (the spin version of Equation 8.4.1.

$$E_\alpha = E_{1s} - \mu_{s,\alpha} \cdot B$$

$$E_\beta = E_{1s} - \mu_{s,\beta} \cdot B$$

Substituting the two equations above into the expression for the photon (Equation 8.4.13) energy gives

$$h\nu = E_\alpha - E_\beta \tag{8.4.14}$$

$$= (E_{1s} - \mu_{s,\alpha} \cdot B) - (E_{1s} - \mu_{s,\beta} \cdot B) \tag{8.4.15}$$

$$= (\mu_{s,\beta} - \mu_{s,\alpha}) \cdot B \tag{8.4.16}$$

Again by analogy with the orbital angular momentum and magnetic moment discussed above, we take the spin magnetic dipole of each spin state, $\mu_{s,\alpha}$ and $\mu_{s,\beta}$, to be related to the total spin angular momentum of each state, S_α and S_β, by a constant **spin gyromagnetic ratio**, γ_s, as shown below.

$$\mu_s = \gamma_s S$$

or each of the two states

$$\mu_{s,\alpha} = \gamma_s S_\alpha$$

$$\mu_{s,\beta} = \gamma_s S_\beta$$

With the magnetic field direction defined as z, the scalar product in Equation 8.4.16 becomes a product of the z-components of the spin angular momenta, $S_{z,\alpha}$ and $S_{z,\beta}$, with the external magnetic field.

Inserting the values for $S_{z,\alpha} = +\frac{1}{2}\hbar$ and $S_{z,\alpha} = -\frac{1}{2}\hbar$ from Equation 8.4.7 and rearranging Equation 8.4.16 yields

$$\frac{h\nu}{B} = -\gamma_s\hbar$$

Calculating the ratio $\frac{h\nu}{B}$ from our experimental results, $\nu = 28 \times 10^9\ Hz$ when $B = 10^4\ gauss$, gives us a value for

$$-\gamma_s\hbar = 18.5464 \times 10^{-21}\ erg/gauss.$$

This value is about twice the Bohr magneton,$-\gamma_e\hbar$, found in Equation 8.4.4 i.e. $\gamma_s\hbar = 2.0023, \gamma_e\hbar$, or

$$\gamma_s = 2.0023\gamma_e \tag{8.4.17}$$

The factor of 2.0023 is called the **g-factor** and accounts for the deviation of the spin gyromagnetic ratio from the value expected for orbital motion of the electron. In other words, it accounts for the spin transition being observed where it is instead of where it would be if the same ratio between magnetic moment and angular momentum held for both orbital and spin motions. The value 2.0023 applies to a freely spinning electron; the coupling of the spin and orbital motion of electrons can produce other values for g.

? Exercise 8.4.1

Carry out the calculations that show that the g-factor for electron spin is 2.0023.

Interestingly, the concept of electron spin and the value g = 2.0023 follow logically from *Dirac's relativistic quantum theory*, which is beyond the scope of this discussion. Electron spin was introduced here as a postulate to explain experimental observations. Scientists often introduce such postulates parallel to developing the theory from which the property is naturally deduced.

8.4: An Electron Has An Intrinsic Spin Angular Momentum is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.4: An Electron has an Intrinsic Spin Angular Momentum** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0. Original source: https://web.archive.org/web/20200619182410/http://www.chemeddl.org.

8.5: Wavefunctions must be Antisymmetric to Interchange of any Two Electrons

Learning Objectives

- Interprete the consequence of exchanging two electrons in a multi-electron atom on the wavelengths
- Connect the Pauli's Exclusion Principle to the permutation symmetry of multi-electron atoms

Quantum mechanics allows us to predict the results of experiments. If we conduct an experiment with indistinguishable particles a correct quantum description cannot allow anything which distinguishes between them. For example, if the wavefunctions of two particles overlap, and we detect a particle, which one is it? The answer to this is not only that we do not know, but that we cannot know. Quantum mechanics can only tell us the probability of finding a particle in a given region. The wavefunction must therefore describe both particles.

The Schrödinger equation for the helium atom is then:

$$\left[-\frac{\hbar^2}{2m}(\nabla_1^2+\nabla_1^2)-V(\mathbf{r}_1)-V(\mathbf{r}_2)+V_{12}(\mathbf{r}_1,\mathbf{r}_2)\right]\psi(\mathbf{r}_1,\mathbf{r}_2)=E\psi(\mathbf{r}_1,\mathbf{r}_2)$$

where the subscripts label each particle, and there are six coordinates, three for each particle. While $|\psi(\mathbf{r}_1,\mathbf{r}_2)\rangle$ is a six dimensional wavefunction (three for each electron) and contains all the information we can measure (a Postulate of Quantum Mechanics), it only provide the probability of finding the electron at a specific volume element, and does **not** tell us which particle is which (e.g., is it electon 1 or electron 2?).

What basis states would be appropriate for $|\psi\rangle$? If we consider the orbital approximation that uses a product wavefunction

$$|\psi(\mathbf{r}_1,\mathbf{r}_2)\rangle=|\varphi_a(\mathbf{r}_1)\rangle|\varphi_b(\mathbf{r}_2)\rangle$$

where $|\varphi_a(\mathbf{r}_1)\rangle$ and $|\varphi_b(\mathbf{r}_2)\rangle$ are one-particle spin-orbitals (with both spin and spatial components) of atoms 1 and 2. This approximation allows us to separate the two particle equation into two one-electron equations:

$$\left[-\frac{\hbar^2}{2m}\nabla_1^2+V(\mathbf{r}_1)\right]|\varphi_a(\mathbf{r}_1)\rangle=E|\varphi_a(\mathbf{r}_1)\rangle$$

$$\left[-\frac{\hbar^2}{2m}\nabla_2^2+V(\mathbf{r}_2)\right]|\varphi_b(\mathbf{r}_2)\rangle=E|\varphi_a(\mathbf{r}_2)\rangle$$

provided that the particles do not interact (e.g., ∇_1^2 does not act on $|\varphi_b(\mathbf{r}_2)\rangle$ and $V_{12}=0$).

Unfortunately, by doing this we have introduced unphysical *labels* to the indistinguishable particles. And this is wrong: the effect of it is that the particles do not interfere with each other because they are in different dimensions (six dimensional space - remember?). When we construct a two particle wavefunction out of two one-particle wavefunctions we must be ensure that the probability density (the measurable quantity $|\psi|^2$) is independent of the **artificial** labels.

The Exchange Operator

We can deepen our understanding of the quantum mechanical description of multi-electron atoms by examining the concepts of electron indistinguishability and the Pauli Exclusion Principle in detail. We will use the following statement as a guide to keep our explorations focused on the development of a clear picture of the multi-electron atom: "When a multi-electron wavefunction is built as a product of single-electron wavefunctions, the corresponding concept is that exactly one electron's worth of charge density is described by each atomic spin-orbital."

A subtle, but important part of the conceptual picture, is that the electrons in a multi-electron system are not distinguishable from one another by any experimental means. Since the electrons are indistinguishable, the probability density we calculate by squaring the modulus of our multi-electron wavefunction also cannot change when the electrons are interchanged (permuted) between different orbitals. In general, if we interchange two identical particles, the world does not change. As we will see below, this requirement leads to the idea that the world can be divided into two types of particles based on their behavior with respect to permutation or interchange.

In order for the probability density to remain unchanged when two particles are permuted, the wavefunction itself can change only by a factor of $e^{i\varphi}$, which represents a complex number, when the particles described by that wavefunction are permuted. As we will show below, the $e^{i\varphi}$ factor is possible because the probability density depends on the absolute square of the function and all expectation values involve $\psi\psi^*$. Consequently $e^{i\varphi}$ disappears in any calculation that relates to the real world because $e^{i\varphi}e^{-i\varphi} = 1$.

We could symbolically write an approximate two-particle wavefunction as $\psi(\mathbf{r}_1, \mathbf{r}_2)$. This could be, for example, a two-electron wavefunction for helium. To exchange the two particles, we simply substitute the coordinates of particle 1 ($\mathbf{r}_1$) for the coordinates of particle 2 ($\mathbf{r}_2$) and vice versa, to get the new wavefunction $\psi(\mathbf{r}_1, \mathbf{r}_2)$. This new wavefunction must have the property that

$$|\psi(\mathbf{r}_1, \mathbf{r}_2)|^2 = \psi(\mathbf{r}_2, \mathbf{r}_1)^*\psi(\mathbf{r}_2, \mathbf{r}_1) = \psi(\mathbf{r}_1, \mathbf{r}_2)^*\psi(\mathbf{r}_1, \mathbf{r}_2) \tag{8.5.1}$$

since the probability density of the electrons in the atom does **not** change upon permutation of the electrons.

? Exercise 8.5.1

Permute the electrons the product function for He wavefunction:

$$\psi(\mathbf{r}_1, \mathbf{r}_2)\rangle = |\varphi_{1s}(\mathbf{r}_1)\rangle|\varphi_{1s}(\mathbf{r}_2)\rangle$$

Equation 8.5.1 will be true only if the wavefunctions before and after permutation are related by a factor of $e^{i\varphi}$,

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = e^{i\varphi}\psi(\mathbf{r}_1, \mathbf{r}_2) \tag{8.5.2}$$

so that

$$\left(e^{-i\varphi}\psi(\mathbf{r}_1, \mathbf{r}_2)^*\right)\left(e^{i\varphi}\psi(\mathbf{r}_1, \mathbf{r}_2)^*\right) = \psi(\mathbf{r}_1, \mathbf{r}_2)^*\psi(\mathbf{r}_1, \mathbf{r}_2) \tag{8.5.3}$$

If we exchange or permute two identical particles twice, we are (by definition) back to the original situation. If each permutation changes the wavefunction by $e^{i\varphi}$, the double permutation must change the wavefunction by $e^{i\varphi}e^{i\varphi}$. Since we then are back to the original state, the effect of the double permutation must equal 1; i.e.,

$$e^{i\varphi}e^{i\varphi} = e^{i2\varphi} = 1 \tag{8.5.4}$$

which is true only if $\varphi = 0$ or an integer multiple of π. The requirement that a double permutation reproduce the original situation limits the acceptable values for $e^{i\varphi}$ to either +1 (when $\varphi = 0$) or -1 (when $\varphi = \pi$). Both possibilities are found in nature.

? Exercise 8.5.2

Use Euler's Equality

$$e^{i\pi} + 1 = 0$$

to show that $e^{i2\varphi} = 1$ when $\varphi = 0$ or $n\pi$ and consequently $e^{i\varphi} = \pm 1$.

We can introduce the *exchange operator* $\hat{P}_{12}$: an operator which permutes the labels of the particles in a multi-particle wavefucntion. This is a rather strange operator, because it only changes the *unphysical* labels which we have attached to the one-particle wavefunctions in order to make the maths more easy. For a meaningful solution we must have a wavefunction which has a probability amplitude unchanged by $\hat{P}_{12}$: it must be either symmetric or antisymmetric with respect to exchange:

$$\hat{P}_{12}|\psi(\mathbf{r}_1, \mathbf{r}_2)\rangle = \pm|\psi(\mathbf{r}_2, \mathbf{r}_1)\rangle \tag{8.5.5}$$

Bosons and femions

Physical solutions must be eigenfunctions of $\hat{P}_{12}$ (i.e., $\hat{H}$ and $\hat{P}_{12}$ commute). Moveover, Equation 8.5.5 argues that the eigenvalues of the Exchange Operator are either +1 (bosons) or −1 (fermions).

Bosons

Wavefunctions for which $e^{i\varphi} = +1$ are defined as symmetric with respect to permutation, because the wavefunction is identical before and after a single permutation. Wavefunctions that are symmetric with respect to interchange of the particles obey the following mathematical relationship,

$$\hat{P}_{12}|\psi(\mathbf{r}_1, \mathbf{r}_2)\rangle = +|\psi(\mathbf{r}_2, \mathbf{r}_1)\rangle$$

The behavior of some particles requires that the wavefunction be symmetric with respect to permutation. These particles are called bosons and have integer spin such as deuterium nuclei, photons, and gluons.

Fermions

The behavior of other particles requires that the wavefunction be antisymmetric with respect to permutation ($e^{i\varphi} = -1$). A wavefunction that is antisymmetric with respect to electron interchange is one whose output changes sign when the electron coordinates are interchanged, as shown below.

$$\hat{P}_{12}|\psi(\mathbf{r}_1, \mathbf{r}_2)\rangle = -|\psi(\mathbf{r}_2, \mathbf{r}_1)\rangle$$

These particles are called fermions and have half-integer spin and include electrons, protons, and neutrinos. Since electrons are fermion, any wavefunction used to describe multiple electrons **must** be antisymmetric with respect to permutation of the electrons. The requirement that the wavefunction be antisymmetric applies to all multi-electron functions $\psi(\mathbf{r}_1, \mathbf{r}_2, \cdots r_i)$, including those approximated as products of single electron functions

$$|\psi(\mathbf{r}_1, \mathbf{r}_2, \cdots r_i)\rangle \approx \varphi_a(\mathbf{r}_1)\varphi_b(\mathbf{r}_2)\cdots\varphi_i(r_i) \tag{8.5.6}$$

? Exercise 8.5.4

What is meant by the term permutation symmetry?

? Exercise 8.5.5

Explain why the product function $\varphi(\mathbf{r}_1)\varphi(\mathbf{r}_2)$ could describe two bosons (deuterium nuclei), but can not describe two fermions (e.g. electrons).

The Pauli's Exclusion Principle

Any physically meaningful Hamiltonian must commute with $\hat{P}_{12}$, otherwise $\hat{H}$ and $\hat{P}_{12}$ could not have common eigenfunctions and the system could not remain in an eigenstate of exchange. A simple product wavefunction like that in Equation 8.5.6 does **not** satisfy this (unless $\varphi_a = \varphi_b$). A **linear combination of all permutations** is required to satisfy indistinguishability constraints.

For a two particle system (e.g., Helium) there is the asymmetric combination

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}|\varphi_a(\mathbf{r}_1)\varphi_b(\mathbf{r}_2) - \varphi_a(\mathbf{r}_2)\varphi_b(\mathbf{r}_1)\rangle \tag{8.5.7}$$

and symmetric combination

$$|\psi^+\rangle = C_{ab}|\varphi_a(\mathbf{r}_1)\varphi_b(\mathbf{r}_2) + \varphi_a(\mathbf{r}_2)\varphi_b(\mathbf{r}_1)\rangle + C_{aa}|\varphi_a(\mathbf{r}_2)\varphi_a(\mathbf{r}_1)\rangle + C_{bb}|\varphi_b(\mathbf{r}_2)\varphi_b(\mathbf{r}_1)\rangle \tag{8.5.8}$$

where the C_{ab} terms are expansion and normalization parameters.

> *Note that the antisymmetric combination (Equation 8.5.7) **cannot** include terms where both particles are in the same state (spin-orbital), but there are three possibilities for the symmetric state (Equation 8.5.8).*

Although any linear combinations of C_{ab}, C_{bb}, and C_{aa} in Equation 8.5.8 is possible, there are three limiting expressions for possible symmetric combinations:

$$|\psi_1^{+i}\rangle = C_{ab}|\varphi_a(\mathbf{r}_1)\varphi_b(\mathbf{r}_2) + \varphi_a(\mathbf{r}_2)\varphi_b(\mathbf{r}_1)\rangle \tag{8.5.9}$$

$$|\psi_2^{+i}\rangle = C_{aa}|\varphi_a(\mathbf{r}_2)\varphi_a(\mathbf{r}_1)\rangle \tag{8.5.10}$$

$$|\psi_3^{+i}\rangle = C_{bb}|\varphi_b(\mathbf{r}_2)\varphi_b(\mathbf{r}_1)\rangle \tag{8.5.11}$$

If $\varphi_a(\mathbf{r}_1) = \varphi_a(\mathbf{r}_2)$, then $|\varphi^{-i}\rangle = 0$. Thus there is no possible antisymmetric combination involving electrons in the same state (spin-orbit). This is the **Pauli exclusion principle**.

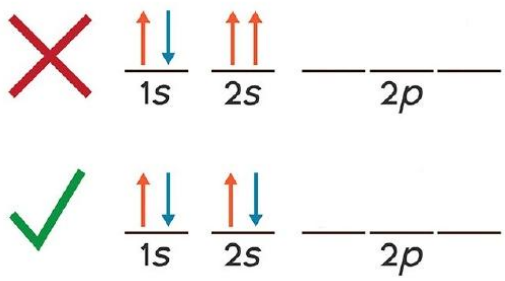


Figure 8.5.1 : The 1s and 2s subshells for beryllium atoms can hold only two electrons and when filled, since each electrons is described by a specific spin-orbital and each orbital (spatial) can hold either an α or a β electron (i.e., they must have opposite spins). Otherwise they will have the same four quantum numbers, in violation of the Pauli Exclusion Principle.

The Pauli Exclusion Principle argues that two electrons **could not** be described by the same spin-orbital. To see the relationship between this statement and the requirement that the wavefunction be antisymmetric for electrons, try to construct an anti-symmetric wavefunction for two electrons that are described by the same spin-orbital

$$|\varphi_b(\mathbf{r}_1)\rangle = \varphi_a(\mathbf{r}_2)\rangle$$

For example, if this were the case for the anti-symmetric combination for helium (Equation 8.5.7), then the wavefunction collapses to zero. We can only constructs wavefunctions that are antisymmetric with respect to permutation symmetry only if each electron is described by a different function.

> *The Pauli Exclusion Principle is simply the requirement that the wavefunction be antisymmetric for electrons, since they are fermions.*

8.5: Wavefunctions must be Antisymmetric to Interchange of any Two Electrons is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.5: Wavefunctions must be Antisymmetric to Interchange of any Two Electrons** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

8.6: Antisymmetric Wave Functions can be Represented by Slater Determinants

Learning Objectives

- Understand how the Pauli Exclusion principle affects the electronic configuration of mulit-electron atoms
- Understand how determinantal wavefunctions (Slater determinents) ensure the proper symmetry to electron permutation required by Pauli Exclusion Principle.
- Connect the electron permutation symmetry requirement to multi-electron wavefunctions to the Aufbau principle taught in general chemistry courses

Let's try to construct an antisymmetric function that describes the two electrons in the ground state of helium. Blindly following the first statement of the Pauli Exclusion Principle, then each electron in a multi-electron atom **must** be described by a different spin-orbital. For the ground-state helium atom, this gives a $1s^2 2s^0 2p^0$ configuration (Figure 8.6.1).

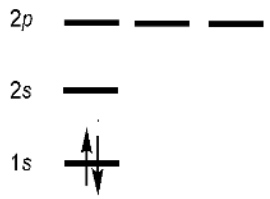


Figure 8.6.1 : Electron configuration for ground state of the helium atom.

We try constructing a simple product wavefunction for helium using two different spin-orbitals. Both have the 1s spatial component, but one has spin function α and the other has spin function β so the product wavefunction matches the form of the ground state electron configuration for He, $1s^2$.

$$|\psi(\mathbf{r}_1, \mathbf{r}_2)\rangle = \varphi_{1s}\alpha(\mathbf{r}_1)\varphi_{1s}\beta(\mathbf{r}_2) \tag{8.6.1}$$

After permutation of the electrons, this becomes

$$|\psi(\mathbf{r}_2, \mathbf{r}_1)\rangle = \varphi_{1s}\alpha(\mathbf{r}_2)\varphi_{1s}\beta(\mathbf{r}_1) \tag{8.6.2}$$

which is different from the starting function since $\varphi_{1s\alpha}$ and $\varphi_{1s\beta}$ are different spin-orbital functions. Hence, the simple product wavefunction in Equation 8.6.1 does not satisfy the indistinguishability requirement since an antisymmetric function must produce the same function multiplied by (–1) after permutation of two electrons, and that is not the case here. We must try something else.

To avoid getting a totally different function when we permute the electrons, we can make **a linear combination of functions**. A very simple way of taking a linear combination involves making a new function by simply adding or subtracting functions. The function that is created by subtracting the right-hand side of Equation 8.6.2 from the right-hand side of Equation 8.6.1 has the desired antisymmetric behavior. The constant on the right-hand side accounts for the fact that the total wavefunction must be normalized.

$$|\psi(\mathbf{r}_1, \mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}}[\varphi_{1s}\alpha(\mathbf{r}_1)\varphi_{1s}\beta(\mathbf{r}_2) - \varphi_{1s}\alpha(\mathbf{r}_2)\varphi_{1s}\beta(\mathbf{r}_1)] \tag{8.6.3}$$

In this orbital approximation, a single electron is held in a single s*pin-orbital* with an orbital component (e.g., the $1s$ orbital) determined by the n, l, m_l quantum numbers and a spin component determined by the m_s quantum number.

The wavefunction in Equation 8.6.3 can be decomposed into spatial and spin components:

$$|\psi(\mathbf{r}_1, \mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}}\underbrace{[\varphi_{1s}(1)\varphi_{1s}(2)]}_{\text{spatial component}}\underbrace{[\alpha(1)\beta(2) - \alpha(2)\beta(1)]}_{\text{spin component}} \tag{8.6.4}$$

Example 8.6.1 : Symmetry to Electron Permutation

Show that the linear combination of spin-orbitals in Equation 8.6.3 is antisymmetric with respect to permutation of the two electrons.

Hint

Replace the minus sign with a plus sign (i.e. take the positive linear combination of the same two functions) and show that the resultant linear combination is symmetric.

Solution

First a reminder of permutation symmetries:

- If the wavefunction is symmetric with respect to permutation of the two electrons then

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

- If the wavefunction is antisymmetric with respect to permutation of the two electrons then

$$|\psi(\mathbf{r}_1,\mathbf{r}_2)\rangle = -|\psi(\mathbf{r}_2,\mathbf{r}_1)\rangle$$

We start with the original wavefunction

$$|\psi(\mathbf{r}_1,\mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}}[\varphi_{1s\alpha}(\mathbf{r}_1)\varphi_{1s\beta}(\mathbf{r}_2) - \varphi_{1s\alpha}(\mathbf{r}_2)\varphi_{1s\beta}(\mathbf{r}_1)]$$

and flip the position of electron 1 with electron 2 and vice versa

$$|\psi(\mathbf{r}_2,\mathbf{r}_1)\rangle = \frac{1}{\sqrt{2}}[\varphi_{1s\alpha}(\mathbf{r}_2)\varphi_{1s\beta}(\mathbf{r}_1) - \varphi_{1s\alpha}(\mathbf{r}_1)\varphi_{1s\beta}(\mathbf{r}_2)] \tag{8.6.5}$$

We then we ask if we can rearrange the left side of Equation 8.6.5 to either become $+|\psi(\mathbf{r}_1,\mathbf{r}_2)\rangle$ (symmetric to permutation) or $-|\psi(\mathbf{r}_1,\mathbf{r}_2)\rangle$ (antisymmetric to permutation).

$$|\psi(\mathbf{r}_2,\mathbf{r}_1)\rangle = \frac{1}{\sqrt{2}}[-\varphi_{1s\alpha}(\mathbf{r}_1)\varphi_{1s\beta}(\mathbf{r}_2) + \varphi_{1s\alpha}(\mathbf{r}_2)\varphi_{1s\beta}(\mathbf{r}_1)]$$

or

$$|\psi(\mathbf{r}_2,\mathbf{r}_1)\rangle = -\frac{1}{\sqrt{2}}[\varphi_{1s\alpha}(\mathbf{r}_1)\varphi_{1s\beta}(\mathbf{r}_2) - \varphi_{1s\alpha}(\mathbf{r}_2)\varphi_{1s\beta}(\mathbf{r}_1)]$$

This is just the negative of the original wavefunction, therefore

$$|\psi(\mathbf{r}_2,\mathbf{r}_1)\rangle = -|\psi(\mathbf{r}_1,\mathbf{r}_2)\rangle$$

The wavefunction is antisymemtric.

Exercise 8.6.1 : Symmetry

Is this linear combination of spin-orbitals

$$|\psi(\mathbf{r}_1,\mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}}[\varphi_{1s\alpha}(\mathbf{r}_1)\varphi_{1s\beta}(\mathbf{r}_2) + \varphi_{1s\alpha}(\mathbf{r}_2)\varphi_{1s\beta}(\mathbf{r}_1)]$$

symmetric or antisymmetric with respect to permutation of the two electrons?

The electronic configuration of the first excited state of He is $1s^1 2s^1 2p^0$ and we can envision four microstates for this configuration (Figure 8.6.2). As spected, the wavefunctions associated for of these microstate must satisfy indistinguishability requirement just like the ground state.

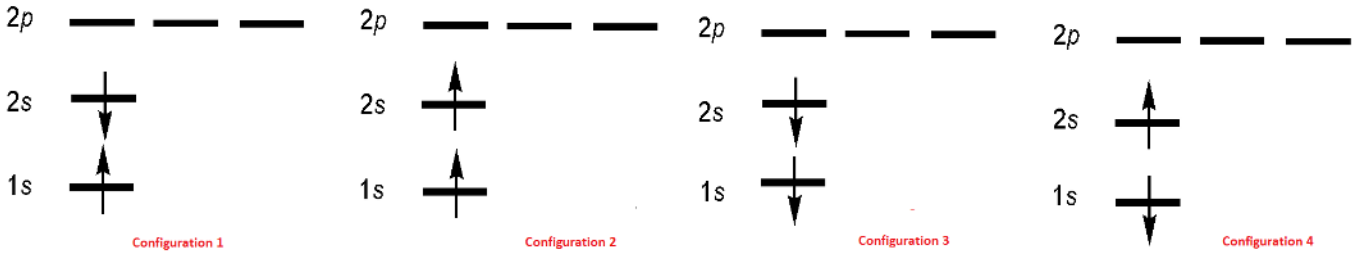


Figure 8.6.2 : Electron configurations for first-excited state of the helium atom.

These electron configurations are used to construct four possible excited-state two-electron wavefunctions (but not necessarily in a one-to-one correspondence):

$$|\psi_1(\mathbf{r}_1,\mathbf{r}_2)\rangle = \frac{1}{2}\underbrace{[\varphi_{1s}(1)\varphi_{2s}(2)+\varphi_{1s}(2)\varphi_{2s}(1)]}_{\text{spatial component}}\underbrace{[\alpha(1)\beta(2)-\alpha(2)\beta(1)]}_{\text{spin component}} \tag{8.6.6}$$

$$|\psi_2(\mathbf{r}_1,\mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}}\underbrace{[\varphi_{1s}(1)\varphi_{2s}(2)-\varphi_{1s}(2)\varphi_{2s}(1)]}_{\text{spatial component}}\underbrace{[\alpha(1)\alpha(2)]}_{\text{spin component}} \tag{8.6.7}$$

$$|\psi_3(\mathbf{r}_1,\mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}}\underbrace{[\varphi_{1s}(1)\varphi_{2s}(2)-\varphi_{1s}(2)\varphi_{2s}(1)]}_{\text{spatial component}}\underbrace{[\alpha(1)\beta(2)+\alpha(2)\beta(1)]}_{\text{spin component}} \tag{8.6.8}$$

$$|\psi_4(\mathbf{r}_1,\mathbf{r}_2)\rangle = \frac{1}{2}\underbrace{[\varphi_{1s}(1)\varphi_{2s}(2)-\varphi_{1s}(2)\varphi_{2s}(1)]}_{\text{spatial component}}\underbrace{[\beta(1)\beta(2)]}_{\text{spin component}} \tag{8.6.9}$$

All four wavefunctions are antisymmetric as required for fermionic wavefunctions (which is left to an exercise). Wavefunctions $|\psi_2\rangle$ and $|\psi_4\rangle$ correspond to the two electrons both having spin up or both having spin down (Configurations 2 and 3 in Figure 8.6.2 , respectively). Wavefunctions $|\psi_1\rangle$ and $|\psi_3\rangle$ are more complicated and are antisymmetric (Configuration 1 - Configuration 4) and symmetric combinations (Configuration 1 + 4). That is, a single electron configuration does not describe the wavefunction. For many electrons, this ad hoc construction procedure would obviously become unwieldy. However, there is an elegant way to construct an antisymmetric wavefunction for a system of N identical particles.

Determinantal Wavefunctions

A linear combination that describes an appropriately antisymmetrized multi-electron wavefunction for any desired orbital configuration is easy to construct for a two-electron system. However, interesting chemical systems usually contain more than two electrons. For these multi-electron systems a relatively simple scheme for constructing an antisymmetric wavefunction from a product of one-electron functions is to write the wavefunction in the form of a determinant. John Slater introduced this idea so the determinant is called a **Slater determinant**.

> John C. Slater introduced the determinants in 1929 as a means of ensuring the antisymmetry of a wavefunction, however the determinantal wavefunction first appeared three years earlier independently in Heisenberg's and Dirac's papers.

The Slater determinant for the two-electron ground-state wavefunction of helium is

$$|\psi(\mathbf{r}_1,\mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}}\begin{vmatrix}\varphi_{1s}(1)\alpha(1) & \varphi_{1s}(1)\beta(1)\\ \varphi_{1s}(2)\alpha(2) & \varphi_{1s}(2)\beta(2)\end{vmatrix} \tag{8.6.10}$$

A shorthand notation for the determinant in Equation 8.6.10 is then

$$|\psi(\mathbf{r}_1,\mathbf{r}_2)\rangle = 2^{-\frac{1}{2}}Det|\varphi_{1s\alpha}(\mathbf{r}_1)\varphi_{1s\beta}(\mathbf{r}_2)| \tag{8.6.11}$$

The determinant is written so the electron coordinate changes in going from one row to the next, and the spin orbital changes in going from one column to the next. The advantage of having this recipe is clear if you try to construct an antisymmetric wavefunction that describes the orbital configuration for uranium! Note that the normalization constant is $(N!)^{-\frac{1}{2}}$ for N electrons.

The generalized Slater determinant for a multi-electrom atom with N electrons is then

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(\mathbf{r}_1) & \varphi_2(\mathbf{r}_1) & \cdots & \varphi_N(\mathbf{r}_1) \\ \varphi_1(\mathbf{r}_2) & \varphi_2(\mathbf{r}_2) & \cdots & \varphi_N(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(\mathbf{r}_N) & \varphi_2(\mathbf{r}_N) & \cdots & \varphi_N(\mathbf{r}_N) \end{vmatrix} \tag{8.6.12}$$

Example 8.6.2 : Helium Atom

Expand the Slater determinant in Equation 8.6.10 for the **He** atom.

Solution

To expand the Slater determinant of the Helium atom, the wavefunction in the form of a two-electron system:

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

This is a simple expansion exercise of a $\mathbf{2 \times 2}$ determinant

$$|\psi(\mathbf{r}_1, \mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}} [\varphi_{1s}(1)\alpha(1)\varphi_{1s}(2)\beta(2) - \varphi_{1s}(2)\alpha(2)\varphi_{1s}(1)\beta(1)]$$

It is not unexpected that the determinant wavefunction in Equation 8.6.10 is the same as the form for the helium wavefunction that is given in Equation 8.6.3.

Exercise 8.6.2 : Lithium Atom

Write and expand the Slater determinant for the ground-state **Li** atom.

Now that we have seen how acceptable multi-electron wavefunctions can be constructed, it is time to revisit the "guide" statement of conceptual understanding with which we began our deeper consideration of electron indistinguishability and the Pauli Exclusion Principle. What does a multi-electron wavefunction constructed by taking specific linear combinations of product wavefunctions mean for our physical picture of the electrons in multi-electron atoms? Overall, the antisymmetrized product function describes the configuration (the orbitals, regions of electron density) for the multi-electron atom. Because of the requirement that electrons be indistinguishable, we cannot visualize specific electrons assigned to specific spin-orbitals. Instead, we construct functions that allow each electron's probability distribution to be dispersed across each spin-orbital. The total charge density described by any one spin-orbital cannot exceed one electron's worth of charge, and each electron in the system is contributing a portion of that charge density.

The four configurations in Figure 8.6.2 for first-excited state of the helium atom can be expressed as the following Slater Determinants

$$|\phi_a(\mathbf{r}_1, \mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_{1s}(1)\alpha(1) & \varphi_{2s}(1)\beta(1) \\ \varphi_{1s}(2)\alpha(2) & \varphi_{2s}(2)\beta(2) \end{vmatrix} \tag{8.6.13}$$

$$|\phi_b(\mathbf{r}_1, \mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_{1s}(1)\alpha(1) & \varphi_{2s}(1)\alpha(1) \\ \varphi_{1s}(2)\alpha(2) & \varphi_{2s}(2)\alpha(2) \end{vmatrix} \tag{8.6.14}$$

$$|\phi_c(\mathbf{r}_1, \mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_{1s}(1)\beta(1) & \varphi_{2s}(1)\alpha(1) \\ \varphi_{1s}(2)\beta(2) & \varphi_{2s}(2)\alpha(2) \end{vmatrix} \tag{8.6.15}$$

$$|\phi_d(\mathbf{r}_1, \mathbf{r}_2)\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_{1s}(1)\beta(1) & \varphi_{2s}(1)\beta(1) \\ \varphi_{1s}(2)\beta(2) & \varphi_{2s}(2)\beta(2) \end{vmatrix} \tag{8.6.16}$$

Slater determinants are constructed by arranging spinorbitals in columns and electron labels in rows and are normalized by dividing by $\sqrt{N!}$, where N is the number of occupied spinorbitals. As you can imagine, the algebra required to compute integrals involving Slater determinants is extremely difficult. It is therefore most important that you realize several things about these states so that you can avoid unnecessary algebra:

- A Slater determinant corresponds to a single electron configuration diagram (Figure 8.6.2). Furthermore, recall that for the excited states of helium we had a problem writing certain stick diagrams as a (space)x(spin) product and had to make linear combinations of certain states to force things to separate (Equation 8.6.7 and 8.6.9). Because of the direct correspondence of configuration diagrams and Slater determinants, the same pitfall arises here: Slater determinants sometimes may not be representable as a (space)x(spin) product, in which case a linear combination of Slater determinants must be used instead. This generally only happens for systems with unpaired electrons (like several of the Helium excited-states).
- A Slater determinant is anti-symmetric upon exchange of any two electrons. We recall that if we take a matrix and interchange two its rows, the determinant changes sign.

The wavefunctions in 8.6.6-8.6.9 can be expressed in term of the four determinants in Equations 8.6.13-8.6.16.

$$\begin{aligned} |\psi_2\rangle &= |\phi_b\rangle \\ &= \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_{1s}(1)\alpha(1) & \varphi_{2s}(1)\alpha(1) \\ \varphi_{1s}(2)\alpha(2) & \varphi_{2s}(2)\alpha(2) \end{vmatrix} \end{aligned}$$

$$\begin{aligned} |\psi_4\rangle &= |\phi_d\rangle \\ &= \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_{1s}(1)\beta(1) & \varphi_{2s}(1)\beta(1) \\ \varphi_{1s}(2)\beta(2) & \varphi_{2s}(2)\beta(2) \end{vmatrix} \end{aligned}$$

but the wavefunctions that represent combinations of spinorbitals and hence combinations of electron configurations (e.g., igure 8.6.2) are combinations of Slater determinants (Equation 8.6.13-8.6.15)

$$\begin{aligned} |\psi_1\rangle &= |\phi_a\rangle - |\phi_c\rangle \\ &= \frac{1}{2} \left(\begin{vmatrix} \varphi_{1s}(1)\alpha(1) & \varphi_{2s}(1)\beta(1) \\ \varphi_{1s}(2)\alpha(2) & \varphi_{2s}(2)\beta(2) \end{vmatrix} - \begin{vmatrix} \varphi_{1s}(1)\beta(1) & \varphi_{2s}(1)\alpha(1) \\ \varphi_{1s}(2)\beta(2) & \varphi_{2s}(2)\alpha(2) \end{vmatrix} \right) \end{aligned}$$

$$\begin{aligned} |\psi_3\rangle &= |\phi_a\rangle + |\phi_c\rangle \\ &= \frac{1}{2} \left(\begin{vmatrix} \varphi_{1s}(1)\alpha(1) & \varphi_{2s}(1)\beta(1) \\ \varphi_{1s}(2)\alpha(2) & \varphi_{2s}(2)\beta(2) \end{vmatrix} + \begin{vmatrix} \varphi_{1s}(1)\beta(1) & \varphi_{2s}(1)\alpha(1) \\ \varphi_{1s}(2)\beta(2) & \varphi_{2s}(2)\alpha(2) \end{vmatrix} \right) \end{aligned}$$

Note the expected change in the normalization constants.

Example 8.6.3 : Carbon Atom

Write the Slater determinant for the ground-state carbon atom. If you expanded this determinant, how many terms would be in the linear combination of functions?

Solution

Carbon has 6 electrons which occupy the 1s 2s and 2p orbitals. Each row in the determinant represents a different electron and each column a unique spin-obital where the electron could be found. There are 6 rows, 1 for each electron, and 6 columns, with the two possible p orbitals both alpha (spin up), in the determinate. There are two columns for each s orbital to account for the alpha and beta spin possibilities. There are two different p orbitals because the electrons in their ground state will be in the different p orbitals and both spin up. N=6 so the normalization constant out front is 1 divided by the square-root of 6!

$$\begin{aligned} &\psi(1,2,3,4,5,6) \\ &= \frac{1}{6!^{1/2}} \begin{vmatrix} \varphi_{1s}(1)\alpha(1) & \varphi_{1s}(1)\beta(1) & \varphi_{2s}(1)\alpha(1) & \varphi_{2s}(1)\beta(1) & \varphi_{2px}(1)\alpha(1) & \varphi_{2py}(1)\alpha(1) \\ \varphi_{1s}(2)\alpha(2) & \varphi_{1s}(2)\beta(2) & \varphi_{2s}(2)\alpha(2) & \varphi_{2s}(2)\beta(2) & \varphi_{2px}(2)\alpha(2) & \varphi_{2py}(2)\alpha(2) \\ \varphi_{1s}(3)\alpha(3) & \varphi_{1s}(3)\beta(3) & \varphi_{2s}(3)\alpha(3) & \varphi_{2s}(3)\beta(3) & \varphi_{2px}(3)\alpha(3) & \varphi_{2py}(3)\alpha(3) \\ \varphi_{1s}(4)\alpha(4) & \varphi_{1s}(4)\beta(4) & \varphi_{2s}(4)\alpha(4) & \varphi_{2s}(4)\beta(4) & \varphi_{2px}(4)\alpha(4) & \varphi_{2py}(4)\alpha(4) \\ \varphi_{1s}(5)\alpha(5) & \varphi_{1s}(5)\beta(5) & \varphi_{2s}(5)\alpha(5) & \varphi_{2s}(5)\beta(5) & \varphi_{2px}(5)\alpha(5) & \varphi_{2py}(5)\alpha(5) \\ \varphi_{1s}(6)\alpha(6) & \varphi_{1s}(6)\beta(6) & \varphi_{2s}(6)\alpha(6) & \varphi_{2s}(6)\beta(6) & \varphi_{2px}(6)\alpha(6) & \varphi_{2py}(6)\alpha(6) \end{vmatrix} \end{aligned}$$

Expanding this determinant would result in a linear combination of functions containing 720 terms. An expanded determinant will contain N! factorial terms, where N is the dimension of the matrix.

8.6.3A Exercise : Excited-State of Helium Atom

Write the Slater determinant for the $1s^1 2s^1$ excited state orbital configuration of the helium atom.

8.6.3B Exercise

Critique the energy level diagram and shorthand electron configuration notation from the perspective of the indistinguishability criterion. Can you imagine a way to represent the wavefunction expressed as a Slater determinant in a schematic or shorthand notation that more accurately represents the electrons? (This is not a solved problem!)

8.6: Antisymmetric Wave Functions can be Represented by Slater Determinants is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.6: Antisymmetric Wavefunctions can be Represented by Slater Determinants** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

8.7: Hartree-Fock Calculations Give Good Agreement with Experimental Data

Learning Objectives

- Understand how the Hartree method is expanded to include symmetrized mulit-electron determential wavefunctions via the Hartree-Fock equations.
- Understand how to calculate the orbital energies from HF theory.
- Apply HF theory with Koopman's theory to estimate ionization energies and electron affinities.

The Hartree method discussed previously is useful as an introduction to the solution of many-particle system and to the concepts of self-consistency and of the self-consistent-field calculations, but its importance is confined to the history of physics. In fact the Hartree method is not just approximate, it is fundamentally wrong since its wavefunction is not antisymmetric to electron permutation! The Hartree-Fock approach discussed below is a better approach, which correctly takes into account the antisymmetric character of the trial wavefunctions.

Although the *Hartree equations* are numerically tractable via the self-consistent field method, it is not surprising that such a crude approximation fails to capture elements of the essential physics. The Pauli exclusion principle demands that the many-body wavefunction be antisymmetric with respect to interchange of any two electron coordinates, e.g.

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = -\Psi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N) \tag{8.7.1}$$

which clearly cannot be satisfied by the multi-electron wavefunctions of the form used in the Hartree Approximation, i.e., the orbital approximation (Equation 8.7.2).

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \approx \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\dots\psi_N(\mathbf{r}_N) \tag{8.7.2}$$

This indistinguishability condition can be satisfied by forming a Slater determinant of single-particle orbitals

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N}} |\psi(\mathbf{r}_1)\psi(\mathbf{r}_2)\dots\psi(\mathbf{r}_N)| \tag{8.7.3}$$

This decouples the electrons resulting in N single-particle **Hartree-Fock** equations:

$$\underbrace{-\frac{\hbar^2}{2m}\nabla^2\psi_i(\mathbf{r})}_{\text{kinetic energy}} + \underbrace{V_{nucleus}(\mathbf{r})\psi_i(\mathbf{r})}_{\text{electron-nucleus potential}} + \underbrace{V_{electron}(\mathbf{r})\psi_i(\mathbf{r})}_{\text{Hartree Term}} - \sum_j \int d\mathbf{r}' \frac{\psi_j^*(\mathbf{r}')\psi_i^*(\mathbf{r}')\psi_j(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} = \epsilon_i\psi_i(\mathbf{r}). \tag{8.7.4}$$

As with the *Hartree equations*, the first term is the kinetic energy of the ith electron

$$-\frac{\hbar^2}{2m}\nabla^2\psi_i(\mathbf{r})$$

and the second term is the electron-nucleus potential between the ith electron and nucleus

$$V_{nucleus}(\mathbf{r})\psi_i(\mathbf{r})$$

The third term (sometimes called the “Hartree” term) is the electrostatic potential between the ith electron and the **average** charge distribution of the other $N-1$ electrons.

$$V_{electron}(\mathbf{r})\psi_i(\mathbf{r}) = J_{j,k} = \int |\phi_j(r)|^2 |\phi_k(r')|^2 \frac{e^2}{r-r'} dr dr' \tag{8.7.5}$$

These three terms are identical to Hartree Equations with the product wavefunction ansatz (i.e., orbital approximation). The fourth term of Equation 8.7.4 is not in the Hartree Equations:

$$\sum_j \int d\mathbf{r}' \frac{\psi_j^*(\mathbf{r}')\psi_i^*(\mathbf{r}')\psi_j(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}$$

and is the *exchange* term. This term resembles the direct Coulomb term, but for the exchanged indices. It is a manifestation of the Pauli exclusion principle, and acts so as to separate electrons of the same spin. This "exchange" term acts only on electrons with the same spin and comes from the Slater determinant form of the wavefunction (Equation 8.7.3). Physically, the effect of exchange is for like-spin electrons to avoid each other. The exchange term adds considerably to the complexity of these equations.

The Hartree-Fock Equations in Equation 8.7.4 can be recast as series of Schrödinger-like equations:

$$\hat{F}|\varphi_i\rangle = \epsilon_i|\varphi_i\rangle \tag{8.7.6}$$

where $\hat{F}$ is called the **Fock operator** and $\{|\varphi_i\rangle\}$ are the Hatree-Fock orbitals with corresponding energies ϵ_i.

The Fock operator is a one-electron operator and solving a Hartree-Fock equation gives the energy and Hartree-Fock orbital for one electron. For a system with 2N electrons, the variable i will range from 1 to N; i.e there will be one equation for each orbital. The reason for this is that only the spatial wavefunctions are used in Equation 8.7.6. Since the spatial portion of an orbital can be used to describe two electrons, each of the energies and wavefunctions found by solving Equation 8.7.6 will be used to describe two electrons.

The nature of the Fock operator reveals how the Hartree-Fock (HF) or Self-Consistent Field (SCF) Method accounts for the electron-electron interaction in atoms and molecules while preserving the idea of **independent atomic orbitals**. The wavefunction written as a Slater determinant of spin-orbitals is necessary to derive the form of the Fock operator, which is

$$\begin{aligned}\hat{F} &= \hat{H}^0 + \sum_{j=1}^{N}(2\hat{J}_j - \hat{K}_j) \\ &= -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r} + \sum_{j=1}^{N}(2\hat{J}_j - \hat{K}_j)\end{aligned} \tag{8.7.7}$$

As shown by the expanded version on the far right, the first term in this equation, $\hat{H}^0$, is the familiar hydrogen-like operator that accounts for the kinetic energy of an electron and the potential energy of this electron interacting with the nucleus. The next term accounts for the potential energy of one electron in an average field created by all the other electrons in the system. The $\hat{J}$ and $\hat{K}$ operators result from the electron-electron repulsion terms in the full Hamiltonian for a multi-electron system. These operators involve the one-electron orbitals as well as the electron-electron interaction energy.

The Fock operator (Equation 8.7.7) depends on **all occupied** orbitals (because of the exchange and Coulomb operators). Therefore, a specific orbital can only be determined if all the others are known. One must use iterative methods to solve the HF equations like the Self-consistent field method discussed previously for the Hartree Approximation.

Exchange Energy

The exchange interaction is a quantum mechanical effect that only occurs between identical particles. Despite sometimes being called an exchange force in analogy to classical force, **it is not a true force**, as it lacks a force carrier. The effect is due to the wavefunction of indistinguishable particles being subject to exchange symmetry, that is, either remaining unchanged (symmetric) or changing its sign (antisymmetric) when two particles are exchanged. Both bosons and fermions can experience the exchange interaction. For fermions, it is sometimes called Pauli repulsion and related to the Pauli exclusion principle. For bosons, the exchange interaction takes the form of an effective attraction that causes identical particles to be found closer together, as in Bose–Einstein condensation.

Example 8.7.1 : Hartree-Fock Energy of Helium

For example, the electron 1 in helium (with $Z = 2$), then

$$\hat{H}^0(1) = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{2e^2}{4\pi\epsilon_0 r_1}$$

The Fock operator is couched in terms of the coordinates of the one electron whose perspective we are taking (which we will call electron 1 throughout the following discussion), and the **average field created by all the other electrons** in the system is built in

terms of the coordinates of a generic “other electron” (which we’ll call electron 2) that is considered to occupy each orbital in turn during the summation over the N spatial orbitals. The best possible one-electron wavefunctions, by definition, will give the **lowest possible total energy** for a multi-electron system used with the complete multielectron Hamiltonian to calculate the expectation value for the total energy of the system. These wavefunctions are called the **Hartree-Fock wavefunctions** and the calculated total energy is the Hartree-Fock energy of the system.

As with the Hartree Equations, solving the Hartree-Fock Equations is mathematically equivalent to assuming each electron interacts only with the average charge cloud of the other electrons. This is how the electron-electron repulsion is handled. This also why this approach is also called the **Self-Consistant Field (SCF)** approach.

Hartree-Fock Energy

The Hartree-Fock equations $h_e\phi_i = \epsilon_i\phi_i$ imply that the orbital energies ϵ_i can be written as:

Multiple \label

where $T+V$ represents the kinetic (T) and nuclear attraction (V) energies, respectively. Thus, ϵ_i is the average value of the kinetic energy plus Coulombic attraction to the nuclei for an electron in ϕ_i plus the sum over all of the spin-orbitals occupied in ψ of Coulomb minus Exchange interactions of these electrons with the electron in ϕ_i.

If ϕ_i is an occupied spin-orbital, the $j=i$ term $[J_{i,i}-K_{i,i}]$ disappears in the above sum and the remaining terms in the sum represent the Coulomb minus exchange interaction of ϕ_i with all of the $N-1$ other occupied spin-orbitals. If ϕ_i is a virtual spin-orbital, this cancelation does not occur because the sum over j does not include $j=i$. So, one obtains the Coulomb minus exchange interaction of ϕ_i with all N of the occupied spin-orbitals in ψ. Hence the energies of occupied orbitals pertain to interactions appropriate to a total of N electrons, while the energies of virtual orbitals pertain to a system with $N+1$ electrons. This difference is very important to understand and to keep in mind.

To give an idea of how well HF theory can predict the ground state energies of several atoms, consider Table 8.7.1 below:

Table 8.7.1 : Hartree-Fock Calculations of Ground Energies of Select Atoms (all energies are in R_H)

Atom	Hartree-Fock Energy	Experiment
He	−5.72	−5.80
Li	−14.86	−14.96
Ne	−257.10	−257.88
Ar	−1053.64	−1055.20

Koopmans' Theorem

Koopmans' theorem states that the first ionization energy is equal to the negative of the orbital energy of the highest occupied molecular orbital. Hence, the ionization energy required to generated a cation and detached electron is represented by the removal of an electron from an orbital without changing the wavefunctions of the other electrons. This is called the "frozen orbital approximation." Let us consider the following model of the detachment or attachment of an electron in an N-electron system.

1. In this model, both the parent molecule and the species generated by adding or removing an electron are treated at the single-determinant level.
2. The Hartree-Fock orbitals of the parent molecule are used to describe both species. It is said that such a model neglects orbital relaxation (i.e., the re-optimization of the spin-orbitals to allow them to become appropriate to the daughter species).

Within this model, the energy difference between the daughter and the parent can be written as follows (ϕ_k represents the particular spin-orbital that is added or removed:

- for electron detachment (vertical ionization energies)

$$E_{N-1} - E_N = -\epsilon_k \tag{8.7.8}$$

- and for electron attachment (electron affinities)

$$E_N - E_{N+1} = -\epsilon_k. \tag{8.7.9}$$

The Hartree-Fock equations deal with exchange exactly; however, the equations neglect more detailed correlations due to many-body interactions. The effects of electronic correlations are not negligible; indeed the failure of Hartree-Fock theory to successfully incorporate correlation leads to one of its most celebrated failures.

Example 8.7.2 : Electron Affinity

Let's derive this result for the case in which an electron is added to the $N+1^{st}$ spin-orbital. The energy of the N-electron determinant with spin-orbitals ϕ_1 through f_N occupied is

$$E_N = \sum_{i=1}^{N} \langle \phi_i | T + V | \phi_i \rangle + \sum_{i=1}^{N} [J_{i,j} - K_{i,j}]$$

which can also be written as

$$E_N = \sum_{i=1}^{N} \langle \phi_i | T + V | \phi_i \rangle + \frac{1}{2} \sum_{i,j=1}^{N} [J_{i,j} - K_{i,j}].$$

Likewise, the energy of the $N+1$-electron determinant wavefunction is

$$E_{N+1} = \sum_{i=1}^{N+1} \langle \phi_i | T + V | \phi_i \rangle + \frac{1}{2} \sum_{i,j=1}^{N+1} [J_{i,j} - K_{i,j}].$$

The difference between these two energies is given by

$$\begin{aligned} E_N - E_{N+1} = & - \langle \phi_{N+1} | T + V | \phi_{N+1} \rangle - \frac{1}{2} \sum_{i=1}^{N+1} [J_{i,N+1} - K_{i,N+1}] \\ & - \frac{1}{2} \sum_{j=1}^{N+1} [J_{N+1,j} - K_{N+1,j}] \\ & = - \langle \phi_{N+1} | T + V | \phi_{N+1} \rangle - \sum_{i=1}^{N+1} [J_{i,N+1} - K_{i,N+1}] \\ & = -\epsilon_{N+1}. \end{aligned}$$

That is, the energy difference is equal to minus the expression for the energy of the $N+1^{st}$ spin-orbital, which was given earlier.

Advanced: Electron Correlation and the "Exchange Hole"

In the Copenhagen Interpretation, the squared modulus of the wavefunction gives the probability of finding a particle in a given place. The many-body wavefunction gives the N-particle distribution function, i.e. $|\Phi(r_1, ..., r_N)|^2$ is the probability density that particle 1 is at r_1, ..., and particle N is at r_N. However, when trying to work out the interaction between electrons, what we want to know is the probability of finding an electron at r, given the positions of all the other electrons $\{r_i\}$. This implies that the electron behaves quantum mechanically when we evaluate its wavefunction, but as a classical point particle when it contributes to the potential seen by the other electrons.

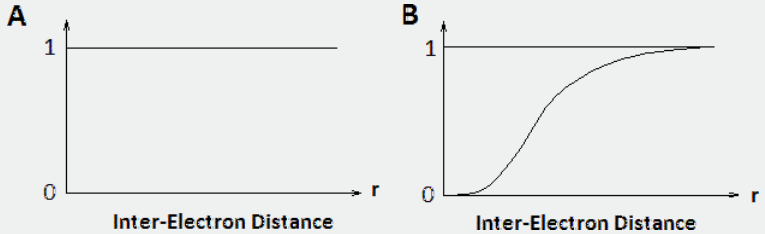


Figure 8.7.2 : Normalized Conditional Probability for from electron-electron interactions (excluding coulomb effects) in (A) the Hartree approximation and (B) the Hartree-Fock approximation.

The contributions of electron-electron interactions in N-electron systems within the Hartree and Hartree-Fock methods are shown in Figure 8.7.2 . The conditional electron probability distributions $n(r)$ of $N-1$ electrons around an electron with given spin situated at $r=0$. Within the Hartree approximation, all electrons are treated as independent, therefore $n(r)$ is structureless. However, within the Hartree-Fock approximation, the N-electron wavefunction reflects the Pauli exclusion principle and near the electron at $r=0$ the exchange hole can be seen where the the density of spins equal to that of the central electron is reduced. Electrons with opposite spins are **unaffected** (not shown).

Summary

So, within the limitations of the HF, frozen-orbital model, the ionization potentials (IPs) and electron affinities (EAs) are given as the negative of the occupied and virtual spin-orbital energies, respectively. This statement is referred to as Koopmans' theorem; it is used extensively in quantum chemical calculations as a means of estimating ionization potentials (Equation 8.7.8) and Electron Affinities (Equation 8.7.9) and often yields results that are qualitatively correct (i.e., ± 0.5 eV). In general Hartree-Fock theory gives a great first order solution (99%) to describing multi-electron systems, but that last 1% is still too great for quantitatively describing many aspects of chemistry and more sophisticated approaches are necessary. These are discussed elsewhere.

8.7: Hartree-Fock Calculations Give Good Agreement with Experimental Data is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.7: Hartree-Fock Calculations Give Good Agreement with Experimental Data** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

8.8: Term Symbols Gives Detailed Descriptions of an Electron Configuration

Learning Objectives

- Understand how electron configurations results in different manifestations of angular momenta (both oribital and spin)
- Describe the manifestations in atoms via atomic term symbols

Atoms have quantum numbers that are directly analogous to the electronic quantum numbers.

The Total Orbital Angular Momentum Quantum Number: L

One might naively think that you could get the total angular momentum of an atom (L) by simply adding up the l values of the individual electrons. The problem with this idea is that the angular momenta of the various electrons are not necessarily pointing in the same direction. Let's consider the case of adding two 2p-orbitals together (i.e., with $l = 1$ quantum numbers).

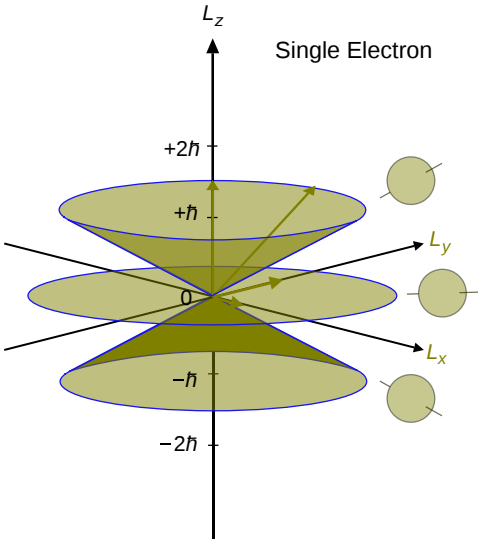


Figure 8.8.1 : The three possible angular momenta vectors for a p-orbital.

As discussed in Section 6.3, each of these electrons has $\sqrt{2}\hbar$ of angular momentum, but oriented in three different directions given by three different m_l values of 1, 0, and -1 (Figure 8.8.1).

$$L = \sum_{i}^{n} l_i \tag{8.8.1}$$

However, we must recognize that while the magnitude of angular momentum is a scalar (represented by l), angular momenta are really vector quantities and to add them together will require vector addition (vs. scalar addition) to do properly and must address all possible orientations of l for each electron.

For example, if two electrons are revolving in the same direction as each other (i.e., same m_l values), one would add just their l values together

$$L(\text{same direction}) = \sqrt{2}\hbar + \sqrt{2}\hbar = 2\hbar$$

If the two elections were revolving in opposites direction (e.., opposite m_l values, you subtract their values.

$$L(\text{same direction}) = \sqrt{2}\hbar - \sqrt{2}\hbar = 0$$

If they are revolving at some off-angle relative to each other (one revolving in a plane and one off plane), you would partially subtract them.

$$L(\text{same direction}) = \sqrt{2}\hbar + 0 = \sqrt{2}\hbar$$

To figure out all of the possible combinations of l for a pair of electrons, simply add them together to get the co-aligned case, subtract them to get the opposing case, and then fill in all the numbers in between to get the off-angle cases. For the two p-orbital case, there are no other possibilities (remember that l and hence L must be non-negative since they represent the magnitude of angular momenta (addressing the m_l values is a different story as we discuss below).

If you prefer to have a formula for the possibilities of L, you can use this:

$$L = |l_1 + l_2|, |l_1 + l_2\text{-}1|, \ldots, |l_1\text{-}l_2| \tag{8.8.2}$$

Technically, Equation 8.8.2 should include all electrons in a system, not just two, to account for all possible combinations of orbital angular momenta, however, there are some tricks that aid in applying this equations to address larger multi-electron atoms. First, electrons in s-orbitals have no orbital angular momentum, so they can be ignored. Second, if block of orbitals are completely filled (e.g., all six p-block spin-orbitals or all 10 d-block spin-orbitals), then the orbital angular momentum vector of each electron will be countered by another electron in the system; these are called **closed-shell** systems. Systems with partially unfilled blocks are called **open-shell** systems.

Hence, conveniently for chemists, an atom's electronic state depends entirely on its unfilled sub shells. Because electrons distribute themselves in a symmetric manner, the inner shell electrons end up canceling out each other's momenta. For an atom in the configuration $1s^2 2s^2 p^2$, only the two p-electrons matter. For an atom in the configuration $1s^2 2s^1 2p^1$, we have to examine only the 2s and 2p electrons (and can ignore the s electrons too).

Example 8.8.1 : Total Orbital Angular Momentum of Carbon

What are the possible L values for the electrons in the $1s^2 2s^2 2p^2$ configuration of carbon?

Solution

Both open-shell electrons (i.e., the 2p electrons) are $l = 1$. The possible combinations are 2, 1, 0.

Example 8.8.2 : Total Orbital Angular Momentum of Unknown Species

What are the possible L values for the electrons in the $[Xe]6s^2 4f^1 5d^1$?

Solution

We can ignore the electrons in the $[Xe]$ core and the electrons in the $6s$ block. So all we have to consider is the f electron $l = 3$ and d electron $l = 2$.

The two extremes possible (Equation 8.8.2) are

$$3 + 2 = 5$$

and

$$3\text{-}2 = 1$$

The possible combinations are thus 5, 4, 3, 2, and 1.

The Total Magnetic Quantum Number: M_l

The Total Magnetic Quantum Number M_l is the total z-component of all of the relevant electrons' orbital momentum. Where L told you how much total angular momentum there is in the system, M_l tells you which direction it is pointing. Like L, a given configuration can have several possible values of M_l, depending on the electrons' relative orientation. Unlike L, M_l is allowed to have negative values. To list the possible M_l values for a two electron system, take the case where both m_l are positive, then take the case where they are both negative, and then fill in the numbers in between (Figure 8.8.1).

$$M_l = m_{l1} + m_{l2}, m_{l1} + m_{l2}\text{-}1, \ldots, \text{-}m_{l1}\text{-}m_{l2} \tag{8.8.3}$$

Note that there is no absolute value function in Equation 8.8.3 like in Equation 8.8.2.

Example 8.8.3 : Total Magnetic Quantum Number of the Zirconium Ground State

What are the possible values M_l of a zirconium atom with the $[Kr]5s^24d^2$ electron configuration?

Solution

Both open-shell electrons (i.e., the 4d electrons) are $l = 2$, so the values are 4, 3, 2, 1, 0, -1, -2, -3, -4.

The Total Spin Magnetic Quantum Number: M_s

M_s is the sum total of the z-components of the electrons' inherent spin. Do not confuse it with M_l, which is the sum total of the z-component of the **orbital angular momentum**. It is easily computed by finding all of the possible combinations of m_s. Since m_s for each individual electron can only be +1/2 or - 1/2, this isn't too complicated.

$$M_s = m_{s1} + m_{s2}, m_{s1} + m_{s2}\text{-}1, ..., m_{s1}\text{-}m_{s2} \tag{8.8.4}$$

Example 8.8.4 : Total Spin Magnetic Quantum Number of the Carbon Ground State

What are the M s values for $1s^22s^22p^2$?

Solution

$M_s = 1, 0, \text{-}1$

The Total Intrinsic Spin Quantum Number: S

The sum total of the spin vectors of all of the electrons is called S. The difference between S and M_s is subtle, but vital for understanding multiplicity. M_s measures the total z-component of the electrons' spins, while S measures the entire resultant vector. The values of S are computing in a manner very similar to M_s. Because S measures the magnitude of a vector, it cannot ever be negative.

$$S = |s_1 + s_2|, |s_1 + s_2\text{-}1|, ..., |s_1\text{-}s_2| \tag{8.8.5}$$

Example 8.8.5 : The Hydrogen Ground State

Find S for $1s^1$.

Solution

S clearly has to be ½ since that's the spin of a single electron and there's only one electron to worry about.

Example 8.8.6 : The Beryllium Excited State

Find S for $1s^22s^12p^1$.

Solution

$S = 1, 0$

Example 8.8.7 : The Carbon Ground State

Find S for carbon atoms with the $1s^22s^22p^2$ electron configuration.

Solution

$S = 1, 0$ This is the same as the previous problem. Notice that S is not affected by which orbitals the electrons are in. S only cares about how many **open-shell electrons** (i.e., unpaired electrons) there are, not about where they are. This is because S measures an inherent property of the electrons themselves

Example 8.8.8

Find S for nitrogen atoms with the $1s^2 2s^2 2p^3$ electron configuration.

Solution

We have not done a three electron case yet, but they are not hard. Find all the combinations for a single pair first, and then factor in the third electron. For two electrons, we already know that the two possible S values are S=1,0. A third electron can either add or subtract ½ from these values, so the final S can be $S = 3/2, 1, 1/2$.

The Total Angular Momentum Quantum Number J

The total orbital angular momentum of an atom (measured in terms of l), and the total spin angular momentum of an atom (measure in S) combine to form total angular momentum, a number that is quantized by the number J. L and S do not necessarily have to be pointing in the same direction (Figure 8.8.1), so J can range from $L + S$ to $|L–S|$.

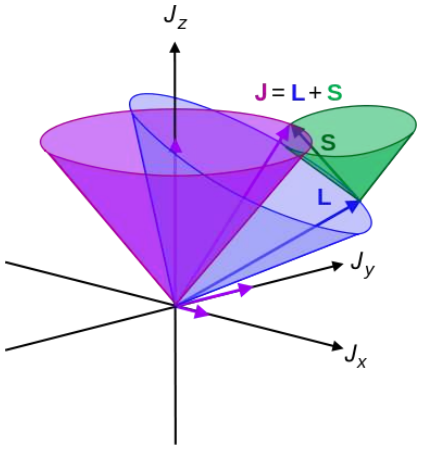


Figure 8.8.1 : Illustration of L-S coupling. Total angular momentum J is purple, orbital L is blue, and spin S is green.

Table 8.8.1 : Quantum Numbers and associated ranges

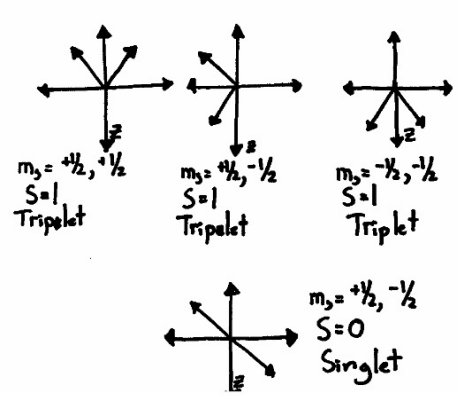
Symbol	Name	Allowed Range
l	Total orbital angular momentum	$\|l_1 + l_2\|, ..., \|l_1 - l_2\|$
M_l	Magnet Quantum number	$[m_{l1} + m_{l2}, ..., -m_{l1} - m_{l2}]$
M_s	Spin Magnetic Quantum Number	$\|m_{s1} + m_{s2}\|, ..., \|m_{s1} - m_{s2}\|$
S	Inherent Spin Number	$\|s_1 + s_2\|, ..., \|s_1 - s_2\|$
M	Multiplicity	2S+1
J	Total Angular Momentum	$L + S, ..., \|L - S\|$

Multiplicity

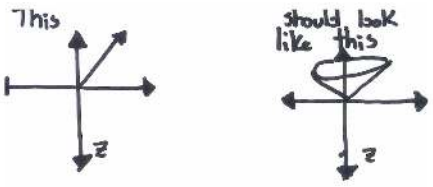
Multiplicity is a simple - sounding concept that defies simple explanations. You know from your first - year education that a singlet is when the net spin (S) is equal to zero (e.g. all the electrons are spin paired), and a triplet happens when the net spin is equal to 1 (e.g. two electrons are pointing in the same direction). They are called “singlet” and “triplet” because there are 3 ways to combine a pair of electron spins to get S=1, but only one way to get $S = 0$. If you draw a picture of the possible ways that two electrons can arrange their spins, you get something like this:



Figure 8.8.1 : When we call an electron spin “up,” what we really mean is that it has a positive z-component (e.g. $m_s = +½$).



While this picture is an improvement over the simple up-down model, it is still misleading. The three spin axes of an electron share a Heisenberg Uncertainty Principle. The more you know about S_x, the less you can know about either S_y or S_z. The same is true for all other combinations of x, y and z. Since we have defined S_z as a known and fixed value, the values of S_x and S_y must be **completely unknown**. This causes the x and y orientations of the electrons to become smeared out across all possible values:



Constructing Term Symbols

Atomic term symbols contain two pieces of information. They tell you the total orbital angular momentum of the atom (l), and they tell you the multiplicity (M). l is denoted by a simple code, similar to the code used to delineate the types of atomic orbitals:

- $L = 0 \rightarrow S$
- $L = 1 \rightarrow P$
- $L = 2 \rightarrow D$
- $L = 3 \rightarrow F$

Note that while the notation is similar, L does NOT say anything about what types of orbitals the electrons are in. A state that has the term symbol P does NOT necessarily have an open p-shell. The multiplicity is indicated by appending a number to the upper left of the symbol. A $L = 2$, $M = 3$ state would be represented by ^3D. The secret to writing the term symbols for an atom is to discover what combinations of l and M are possible for that atom with that specific electronic configuration. An atom that only has closed shells will always be $1S$.

Each term symbol represents a discrete energy level. We can place these levels in the correct order by using these simple rules:

- 1: High multiplicity values mean low energy
- 2: If there is a tie, high l values mean low energy
- 3a: If there is still a tie and the shell is less than half full, then low J means low energy
- 3b: If the shell is more than half full, then high J means low energy

These rules reliably predict the ground state. They have only erratic agreement with experiment when ordering the other levels.

Example 8.8.9 : Hydrogen Ground State

What are the term symbols for the microstates possible for $1s^1$ electronic configuration of hydrogen?

Solution

Since there is only one electron, this is a simple problem. $L = 0$ and $M = 1$, so the only possible term symbol is ^2S. With only one electron, $S = ½$, so $J = 0 + ½ = ½$. Only one microstate exists for this configuration and it has a term symbol of $^2S_{½}$.

Example 8.8.10 : Boron

What are the term symbols for the microstates possible for $1s^2 2s^2 2p^1$ electronic configuration of boron?

Solution

There still only one open shell electron, so $L = 1$, $M = 1$ and $S = ½$. We get a term symbol of the type ^2P, which gets split into separate symbols because $J = 3/2$ and $1/2$. Two possible microstates exist for this system with term symbols of $^2P_{3/2}$ and $^2P_{1/2}$.

Example 8.8.11 : Beryllium Excited State

What are the term symbols for the microstates possible for the $1s^2 2s^1 2p^1$ excited-state electronic configuration of Beryllium?

Solution

Now we have two electrons to worry about. Since $l_1 = 0$ and $l_2 = 1$, the only possible combination is $L = 1$. The possible combinations of S are: $S = 1, 0$. This means that $M = 3$ or $M = 1$. The term symbols will be of the form ^1P and ^3P. For the ^1P state, $L = 1$ and $S = 0$, so $J = 1$. For the second state, $L = 1$ and $S = 1$, so $J = 2, 1, 0$. There are four microstates for this configuration with term symbols of $^1P_1$ and $^3P_2$, $^3P_1$, and $^3P_0$.

Example 8.8.12 : Zirconium

What are the term symbols for the microstates possible for the $[Kr]5s^2 4d^2$ ground-state electronic configuration of zirconium?

Solution

This is a much harder problem. We will need to use a special technique to disentangle all of the possible combinations of l and M. Let's start be listing the relevant quantum numbers for the two open-shell electrons:

$l_1 = 2$	$l_2 = 2$
$m_{l1} = 2, 1, 0, -1, -2$	$m_{l2} = 2, 1, 0, -1, -2$
$m_{s1} = ½, -½$	$m_{s1} = ½, -½$

Let's combine these numbers to generate the atomic quantum numbers:

$$L = 4, 3, 2, 1, 0$$

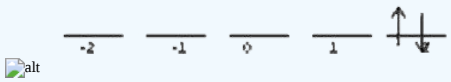
$$M_l = 4, 3, 2, 1, 0, -1, -2, -3, -4$$

$$M_s = 1, 0$$

We know that there will at least one each of S, P, D, F and G. It isn't immediately clear which of these will be singlets and which will be triplets. To figure this out, we need to systematically examine the possible microstates. It turns out that there are **45** possible ways to put distribute two electrons between 5 d orbitals. That's a lot! The easiest way to list the states is to organize them into a chart:

$M_s = -1$	$M_s = 0$	$M_s = +1$

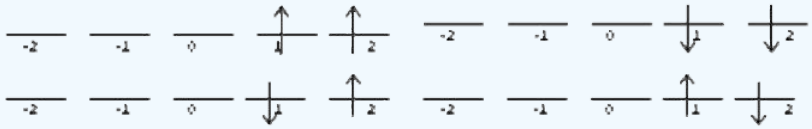
Attacking the chart one row at a time. Ask yourself, how many ways can I arrange the two electrons to give me $M_l = 4$? It turns out there is only one possible combination that does this:



This state is has $M_s = 0$. This means that there is only 1 microstate that corresponds to $M_l = 4$ and $M_s = 0$, and none that correspond to $M_l = 4$ and $M_s = \pm 1$. We add this microstate to the chart like this:

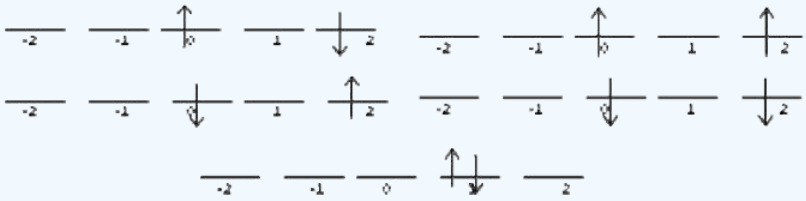
M_l	$M_s = -1$	$M_s = 0$	$M_s = +1$
4	0	1	0

Now, how many ways are there to get $M_l = 3$?



M_l	$M_s = -1$	$M_s = 0$	$M_s = +1$
4	0	**1**	0
3	1	2	1

For $M_l = 2$, we find the following states:



M_l	$M_s = -1$	$M_s = 0$	$M_s = +1$
4	0	**1**	0
3	1	2	1
2	1	3	1

You should be able to draw the microstates on your own by now. You should find 8 states, four of which are singlet and four of which are triplets.

M_l	$M_s = -1$	$M_s = 0$	$M_s = +1$
4	0	**1**	0
3	1	2	1
2	1	3	1
1	2	4	2

There are only nine possible ways to arrange the electrons to get $M_l = 0$

M_l	$M_s = -1$	$M_s = 0$	$M_s = +1$
4	0	**1**	0
3	1	2	1
2	1	3	1
1	2	4	2
0	2	5	2

The rest of the chart will be symmetric to the first half, so we do not need to do any more work:

M_l	$M_s = -1$	$M_s = 0$	$M_s = +1$
4	0	1	0
3	1	2	1
2	1	3	1
1	2	4	2
0	2	5	2
-1	2	4	2
-2	1	3	1
-3	1	2	1
-4	0	1	0

Now that we have a listing of all of the microstates, we need to figure out how to divide them up between the term symbols. It turns out that each term symbol can have, at most, one microstate from each box on the chart. The term symbols always end up claiming a “box” of microstates, centered on the middle of the chart. This is easier shown than said.

Attacking the chart from the top, we can see that the $M_l = 4\ M_s = 0$ state clearly belongs to a ^1G symbol. The $M_l = -4\ M_s = 0$ box also clearly belongs to this symbol. If I connect these states with a "box," I get this:

M_l	$M_s = -1$	$M_s = 0$	$M_s = +1$
~~4~~	~~0~~	~~1~~	~~0~~
3	1	2	1
2	1	3	1
1	2	4	2
0	2	5	2
-1	2	4	2
-2	1	3	1
-3	1	2	1
~~-4~~	~~0~~	~~1~~	~~0~~

The strickthough configurations all belong to the ^1G state. Let's subtract them from the chart to indicate that they are not available for other term symbols.

M_l	$M_s = -1$	$M_s = 0$	$M_s = +1$
3	1	2	1
2	1	3	1
1	2	4	2
0	2	5	2
-1	2	4	2
-2	1	3	1
-3	1	2	1

The next row indicates a $L = 3$ state. Because there are three M_s values available, this is a **triplet**. The term symbol will be ^3F, which reduces the chart down to

M_l	$M_s = -1$	$M_s = 0$	$M_s = +1$
1	2	4	2
0	2	5	2
-1	2	4	2

The next state will be ^1D. Leaving us with

M_l	$M_s = -1$	$M_s = 0$	$M_s = +1$
1	1	1	1
0	1	2	1
-1	1	1	1

Next is a ^3P state. The chart is getting pretty small now wtih

M_l	$M_s = 0$
	0

The last remaining microstate comprises the ^1S term symbol.

The total listing is

$$^1G, \, ^3F, \, ^1D, \, ^3P, \, ^1S.$$

Assigning J values, we get

$$^1G_4, \, ^3F_4, \, ^3F_3, \, ^3F_2, \, ^1D_2, \, ^3P_2, \, ^3P_1, \, ^3P_0, \, ^1S_0$$

If you can do this problem, you can do almost any atomic term symbol.

Note

The secret to writing the term symbols for an atom is to discover what combinations of l and M are possible for that atom with that specific electronic configuration.

Shortcuts

There is a deep symmetry that connects different electronic configurations. It turns out that a p^1 configuration has the same term symbols as a p^5. Similarly, $p^2 = p^4$. A similar relationship can be used to figure out high electron number term symbols for the d and f orbitals.

8.8: Term Symbols Gives Detailed Descriptions of an Electron Configuration is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.8: Term Symbols Gives a Detailed Description of an Electron Configuration** by Mattanjah de Vries is licensed CC BY 4.0.

8.8B: Multi-electron Considerations - A Closer Look at Helium

If we have two identical fermions of spin $1/2$, confined in the same region, what is the appropriate wavefunction? In the scattering case we could measure spins far from the interaction, and if we knew that the total spins is conserved, spins can be associated with each particle. In the bound state we cannot tell which particle we are measuring, so the ket must contain both spin and spatial wavefunctions of both particles.

Assuming the spins do not interact, we can separate the two-particle spin-only (ignoring the spatial part) wavefunction into $\sigma(1,2) = \sigma_1\sigma_2$. We also know the appropriate one particle basis states $\uparrow 1$, $\downarrow 1$, $\uparrow 2$, $\downarrow 2$, where $\uparrow 1$ represents "particle 1" in spin up. The ONLY combinations for **indistinguishable particles** are then:

$$\uparrow 1 \uparrow 2 \tag{8.8B.1}$$

$$\downarrow 1 \downarrow 2 \tag{8.8B.2}$$

$$\frac{1}{\sqrt{2}}(\uparrow 1 \downarrow 2 + \downarrow 1 \uparrow 2) \tag{8.8B.3}$$

$$\frac{1}{\sqrt{2}}(\uparrow 1 \downarrow 2 - \downarrow 1 \uparrow 2) \tag{8.8B.4}$$

- Operating these four possible spin-only wavefunctions with $\hat{P}_{12}$ yields the eigenvalues: 1, 1, 1 and -1 respectively.
- $S^2 = S(S+1)$ on yields: 2, 2, 2 and 0,
- S_z yields: 1, -1, 0 and 0.

Thus the demands of indistinguishability couples the spins of two identical particles into a triplet ($S = 1$) and a singlet ($S = 0$). The spin-1 vector (i.e., $S = 1$) has **three** possible M_s component values (wavefunctions 8.8B.1, 8.8B.2, and 8.8B.3). This is reflected in its $2S+1$ multiplicity and is why it is called a triplet state.

The *overall* wavefunction describing fermions must be antisymmetric with respect to exchange, i.e.

$$\hat{P}_{12}|\psi\rangle = -|\psi\rangle \tag{8.8B.5}$$

Therefore in an atom or molecule where Φ includes both spin and spatial parts, the spin and spatial parts of a fermionic wavefunction must have opposite exchange symmetry. Spin must be considered even if the energy (Coulomb potential) depends explicitly only on the spatial part. As discussed in the next section, the expectation value of the potential energy is different for symmetric and antisymmetric spatial combinations.

The Exchange Interaction

Since the *overall* wavefunction describing fermions must be antisymmetric with respect to exchange, i.e.

$$\hat{P}_{12}|\psi\rangle = -|\psi\rangle, \tag{8.8B.6}$$

therefore in an atom or molecule where Φ includes both spin and spatial parts, the spin and spatial parts of a fermionic wavefunction must have opposite exchange symmetry. Spin must be considered even if the energy (Coulomb potential) depends explicitly only on the spatial part. However, the expectation value of the potential energy is different for symmetric and antisymmetric spatial combinations.

For helium, there asymmetric combination is

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}|\phi_a(r_1)\phi_b(r_2) - \phi_a(r_2)\phi_b(r_1)\rangle \tag{8.8B.7}$$

and symmetric combination is

$$|\Phi^+\rangle = C_{ab}|\phi_a(r_1)\phi_b(r_2) + \phi_a(r_2)\phi_b(r_1)\rangle + C_{aa}|\phi_a(r_2)\phi_a(r_1)\rangle + C_{bb}|\phi_b(r_2)\phi_b(r_1)\rangle \tag{8.8B.8}$$

For convenience, let's consider only the $|\Phi^{\pm}\rangle$ combinations with $C_{aa} = C_{bb} = 1$). The corresponding energy associated with the inter-electron interaction potential V can be expanded

$$\langle \psi^{\pm}|\hat{V}|\psi^{\mp}\rangle = \langle \phi_a(r_1)\phi_b(r_2)|V(r)|\phi_a(r_1)\phi_b(r_2)\rangle \pm \langle \phi_a(r_1)\phi_b(r_2)|V(r)|\phi_a(r_2)\phi_b(r_1)\rangle \tag{8.8B.9}$$

The first term is called the **Coulomb interaction (J)** and the second term is known as **the exchange interaction (\K\))**: a measurable contribution to the energy comparable in size to the first, which has no classical analog.

Now notice something strange. The exchange interaction has split the $S = 1$ states from the $S = 0$ states. We could write the potential as

$$\hat{V} = J_{nl} - (2\hat{S} - 1)K_{nl} \tag{8.8B.10}$$

even though the Hamiltonian does not act on the spin! This is because the sign of the exchange integral depends on the (anti)symmetry of the spatial wavefunction. Thus we can write the matrix element as

$$\langle \psi | J_{nl} - (2S - 1)K_{nl}|\psi\rangle \tag{8.8B.11}$$

> *This 'exchange interaction' depends on the spin; the triplet states have lower energy than the singlet (this is one of Hund's rules for determining energy levels in atoms).*

Helium

Helium is the simplest system for which we are unable to accurately calculate the energy. For a single electron moving in the field of a helium nucleus, the spatial wavefunctions are similar to those of hydrogen $|\phi_{n,l,m_s}\rangle$. When a second electron is added, a reasonable basis set is exchange-symmetrized wavefunctions consisting of spin states multiplying hydrogenic spatial parts:

$$(\phi_{n,l,m}(r_1)\phi_{n',l',m'_l}(r_2) \pm \phi_{n,l,m}(r_2)\phi_{n',l',m'_l}(r_1)) \tag{8.8B.12}$$

The **overall wavefunction** must be antisymmetric so:

- the singlet spin states (exchange-antisymmetric) must combine with symmetric spatial states and
- the triplet spin states (exchange-symmetric) must combine with antisymmetric spatial states.

If both electrons were in the **same spatial state (orbital)**, the antisymmetric spatial wavefunction would be:

$$|(a(r_1)a(r_2) - a(r_2)a(r-1))\rangle = 0 \tag{8.8B.13}$$

Hence, there is no **triplet** for the ground state.

Electron-electron interaction - ground state by perturbation theory

The hydrogen wavefunctions are only a choice of basis set: the hydrogenic potential ignores the electron-electron repulsion. A simple approach is to treat this as a perturbation and to use degenerate perturbation theory. The perturbation potential is just

$$V = \frac{e^2}{4\pi\epsilon_0 r_{12}} \tag{8.8B.14}$$

where $r_{12} = |r_1 - r_2|$. For helium, the **unperturbed** spatial ground state is just a product of the $1s$ hydrogenic ones with $Z = 2$:

$$\phi_{n=1,l=0,m_l=0}(r_1)\phi_{n=1,l=0,m_l=0}(r_2) = \frac{Z^3}{\pi a_0^3}\exp[-Zr_1/a_0]\exp[-Zr_2/a_0] \tag{8.8B.15}$$

The zero-order energy (i.e., unperturbed energy) associated with this wavefunction is

$$E^0 = \frac{4Ze^4}{\hbar^2} \tag{8.8B.16}$$

Via first-order perturbation theory, the energy shift from the perturbed energy of any state is

$$\langle \phi | \frac{e^2}{4\pi\epsilon_0 r_{12}} | \phi \rangle \tag{8.8B.17}$$

and for the ground state fo Helium, Equation 8.8B.17 is given by:

$$\langle \phi_{n=1,l=0,m_l=0}(r_1)\phi_{n=1,l=0,m_l=0}(r_2)| \frac{e^2}{4\pi\epsilon_0 r_{12}} |\phi_{n=1,l=0,m_l=0}(r_1)\phi_{n=1,l=0,m_l=0}(r_2)\rangle \tag{8.8B.18}$$

Comparison of the zero-order energy and the measured energy show that electron-electron repulsion is over 30% of the unperturbed energy (Equation 8.8B.16), so perturbation theory may seem inappropriate (although higher order perturbation theory may be used). Strictly, Equation 8.8B.18 is not even the right integral, as it neglects electron correlation effects (discussed elsewhere), but the value of this integral

$$\frac{5Ze^4}{8\hbar^2} \tag{8.8B.19}$$

is within 5% of the actual energy.

> *Note also that the radial wavefunctions are different for* $2s$ *and* $2p$*, so the electron-electron interaction splits the degeneracy between* $1s2s$ *and* $1s2p$ *configurations.*

Multiplicity and Degeneracy of Excited States

Ignoring electron-electron interaction (e.g., in the hydrogen atom), all $1s2s$ and $1s2p$ states have the same energy. However, the electron-electron perturbation

$$H^1 = \frac{e^2}{4\pi\epsilon_0 r_{12}} \tag{8.8B.20}$$

lifts that degeneracy, which we can treat with **degenerate** perturbation theory (this is outside the scope of this class). Rather than evaluating resulting perturbation theory integral exactly, we can use a physical argument about the nature of the perturbation. Since $(e^2/4\pi\epsilon_0 r_{12})$ is **not** an externally applied potential, it applies no net torque nor force on the electrons. The perturbation cannot change the angular momentum, so it cannot mix states with different l or m_l (i.e., the perturbed wavefunctions does not have components of differing angular momentum quantum numbers). Integrating over θ results in $\delta_{ll'}$ and integrating over ϕ give $\delta_{m_l m_l'}$, so **total orbital angular momentum** remains a good quantum number, hence:

$$L = \sum_{i=1}^{2} l_i = 0 + 0 = 0 \tag{8.8B.21}$$

for the $1s^1 2s^1$ electron configuration (promoting a 1s electron to the 2s orbital) or

$$L = \sum_{i=1}^{2} l_i = 0 + 1 = 1 \tag{8.8B.22}$$

for the $1s^1 2p^1$ electron configuration (promoting a 1s electron to the 2p orbital).

> *Different electron configurations will have different total orbital angular momentum and spin angular momentum amplitudes.*

Since the 2s electron is less well screened from the nuclear charge by the $1s$ that the $2p$ electron due to its closer radial distribution function, it will have lower energy (see Slater's Rules). As wth the helium ground-state, for a given spatial excited state, the possible normalized spin-orbital wavefunction combinations, consistent with the antisymmetry requirement are a triplet ($S = 1$) and a singlet ($S = 0$).

The spin-orbital wavefunctions for the triplet state are

$$\Phi_3 = \frac{1}{\sqrt{2}}[\phi_{n,l,m_l,n',l',m_l'} - \phi_{n',l',m_l',n,l,m_l}](\uparrow\uparrow) \tag{8.8B.23}$$

and

$$\frac{1}{\sqrt{2}}[\phi_{n,l,m_l,n',l',m_l'} - \phi_{n',l',m_l',n,l,m_l}](\downarrow\downarrow) \tag{8.8B.24}$$

and

$$\frac{1}{2}[\phi_{n,l,m_l,n',l',m_l'} - \phi_{n',l',m_l',n,l,m_l}](\downarrow\uparrow + \downarrow\uparrow) \tag{8.8B.25}$$

The spin-orbital wavefunction for the singlet state is

$$\Phi_1 = \frac{1}{2}[\phi_{n,l,m_l,n',l',m_l'} + \phi_{n',l',m_l',n,l,m_l}](\downarrow\uparrow - \downarrow\uparrow) \tag{8.8B.26}$$

where $\phi_{n,l,m_l,n',l',m_l'}\rangle$ represents electron 1 in a hydrogenic state with quantum numbers n, l and m_l and electron 2 with n', l', and m_l. The subscripts on the Φ label spin *multiplicity*

$$2S + 1 \tag{8.8B.27}$$

Again, whole effect of the potential is contained in the spatial part, the spin integral will be $\delta_{\sigma,\sigma'}$. So off-diagonal matrix elements are all zero. We need to evaluate two types of integrals to compute the energies of these states:

- the **Coulomb integral**

$$J_{nl} = \langle\phi_{n,l,m_l,n',l',m_l'}|\frac{e^2}{4\pi\epsilon_0 r_{12}}|\phi_{n,l,m_l,n',l',m_l'}\rangle \tag{8.8B.28}$$

- and the **Exchange integral**

$$K_{nl} = \langle\phi_{n,l,m_l,n',l',m_l'}|\frac{e^2}{4\pi\epsilon_0 r_{12}}|\phi_{n',l',m_l',n,l,m_l}\rangle \tag{8.8B.29}$$

Within first-order perturbation theory, an energy shift can be calculated by Equation 8.8B.17. Apply this expression to the $1s^1 2s^1$ singlet state wavefunction (Equation 8.8B.26) results in:

$$\begin{aligned} E^1 = \frac{1}{2}\frac{e^2}{4\pi\epsilon_o}\Big(&\langle\phi_{100,200}|\frac{1}{r_{12}}|\phi_{100,200}\rangle + \langle\phi_{200,100}|\frac{1}{r_{12}}|\phi_{200,100}\rangle \\ &\pm \langle\phi_{100,200}|\frac{1}{r_{12}}|\phi_{200,100}\rangle \pm \langle\phi_{200,100}|\frac{1}{r_{12}}|\phi_{100,200}\rangle\Big) \end{aligned} \tag{8.8B.30}$$

which in terms of the integrals defined above can be rewritten

$$E^1 = \frac{1}{2}\left(J_{1,2} + J_{2,1} \pm K_{1,2} \pm K_{2,1}\right) \tag{8.8B.31}$$

since $J_{1,2} = J_{2,1}$ and $K_{1,2} = k_{2,1}$, then

$$E^1 = J \pm K \tag{8.8B.32}$$

where the + applies to the singlet state (the 1s and 2s electrons have opposing spin) and the − to the triplet (the electrons aligned). The Coulomb integral as a measure of the electron-electron repulsion, increases the energy of the state. However, the exchange integral can either increase (for singlets) or decrease energy (for triplets). This means that electron-electron repulsion lifts the degeneracy two-fold:

- the energy levels are split by different direct interactions into $L = 0$ and $L = 1$ and
- the energy levles are again split through exchange interaction into singlet and triplet.

The final degeneracies of states of helium with one electron excited to $n = 2$ are 3, 1, 9 and 3 (Figure 8.8B.1).

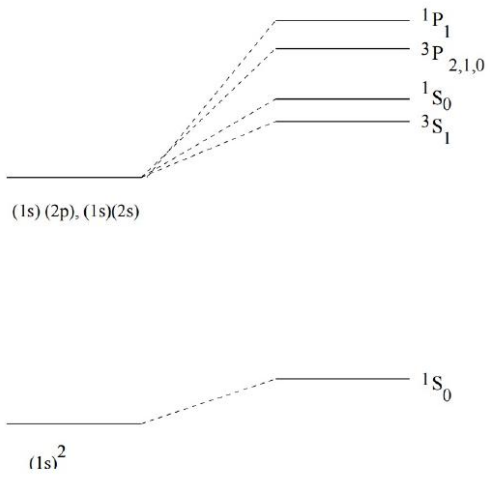


Figure* 8.8*B*.1*: *Ground and Excited-states of Helium. (left) underperturned energies from* H^0 *and (right) perturbed energies due to electron-electron repulsion.*

The spectroscopic notation in Figure 8.8*B*.1 is called a term symbol and expresses the quantum numbers as:

$$^{2S+1}L_J \tag{8.8B.33}$$

where J is the total angular moment.

> *The 'exchange force' selects preferred spin state via the requirement of overall antisymmetry of the wavefunction.*

Contributors

- Graeme Ackland (University of Edinburgh)

8.8B: Multi-electron Considerations - A Closer Look at Helium is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

8.9: The Allowed Values of J - the Total Angular Momentum Quantum Number

Learning Objectives

- Compate two spin-orbit coupling schemes that couple the total spin angular momenta and total orbital angular momenta of a multi-electron spectra

We need to be able to identify the electronic states that result from a given electron configuration and determine their relative energies. An electronic state of an atom is characterized by a specific energy, wavefunction (including spin), electron configuration, total angular momentum, and the way the orbital and spin angular momenta of the different electrons are coupled together. There are two descriptions for the coupling of angular momentum. One is called **j-j coupling**, and the other is called **L-S coupling**. The j-j coupling scheme is used for heavy elements (z > 40) and the L-S coupling scheme is used for the lighter elements. Only L-S coupling is discussed below.

L-S Coupling of Angular Momenta

L-S coupling also is called R-S or Russell-Saunders coupling. In L-S coupling, the orbital and spin angular momenta of all the electrons are combined separately

$$L = \sum_i l_i \tag{8.9.1}$$

$$S = \sum_i s_i \tag{8.9.2}$$

The total angular momentum vector then is the sum of the total orbital angular momentum vector and the total spin angular momentum vector.

$$J = L + S \tag{8.9.3}$$

The total angular momentum quantum number parameterizes the total angular momentum of a given particle, by combining its orbital angular momentum and its intrinsic angular momentum (i.e., its spin). Due to the spin-orbit interaction in the atom, the orbital angular momentum no longer commutes with the Hamiltonian, nor does the spin.

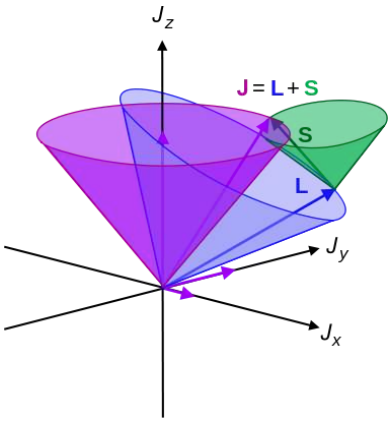


Figure 8.9.1 : "Vector cones" of total angular momentum $\boldsymbol{J}$ (purple), orbital $\boldsymbol{L}$ (blue), and spin $\boldsymbol{S}$ (green). The cones arise due to quantum uncertainty between measuring angular momentum components (see vector model of the atom). (Public Domain; Maschen).

However the total angular momentum $\boldsymbol{J}$ does commute with the Hamiltonian and so is a constant of motion (does not change in time). The relevant definitions of the angular momenta are:

Orbital Angular Momentum

$$|\vec{L}| = \hbar\sqrt{\ell(\ell+1)}$$

with its projection on the z-axis

$$L_z = m_\ell \hbar$$

Spin Angular Momentum

$$|\vec{S}| = \hbar\sqrt{s(s+1)}$$

with its projection on the z-axis

$$S_z = m_s \hbar$$

Total Angular Momentum

$$|\vec{J}| = \hbar\sqrt{j(j+1)}$$

with its projection on the z-axis

$$J_z = m_j \hbar$$

where

- l is the azimuthal quantum number of a single electron,
- s is the spin quantum number intrinsic to the electron,
- j is the total angular momentum quantum number of the electron,

The quantum numbers take the values:

$$m_\ell \in \{-\ell, -(\ell-1) \cdots \ell-1, \ell\}, \quad \ell \in \{0, 1 \cdots n-1\} \tag{8.9.4}$$
$$m_s \in \{-s, -(s-1) \cdots s-1, s\}, \tag{8.9.5}$$
$$m_j \in \{-j, -(j-1) \cdots j-1, j\}, \tag{8.9.6}$$
$$m_j = m_\ell + m_s, \quad j = |\ell + s| \tag{8.9.7}$$

and the magnitudes are:

$$|\mathbf{J}| = \hbar\sqrt{j(j+1)} \tag{8.9.8}$$
$$|\mathbf{J_1}| = \hbar\sqrt{j_1(j_1+1)} \tag{8.9.9}$$
$$|\mathbf{J_2}| = \hbar\sqrt{j_2(j_2+1)} \tag{8.9.10}$$

in which

$$j \in \{|j_1 - j_2|, |j_1 - j_2| - 1 \cdots j_1 + j_2 - 1, j_1 + j_2\}$$

This process may be repeated for a third electron, then the fourth etc. until the total angular momentum has been found.

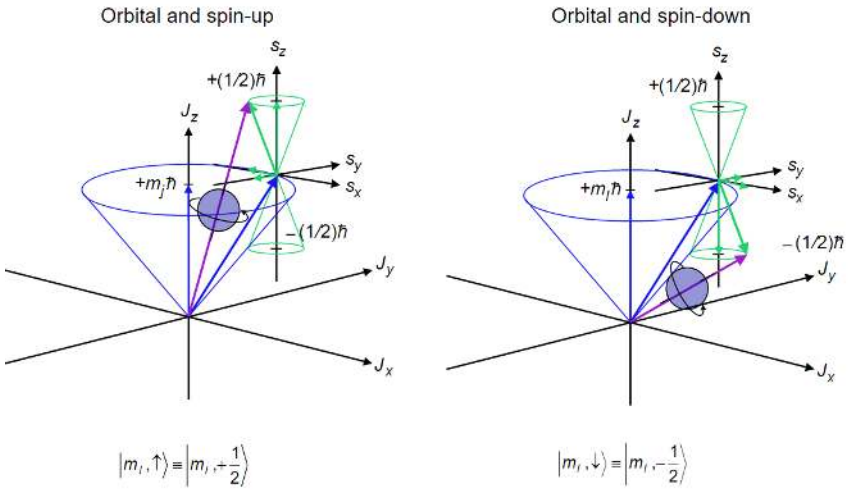


Figure 8.9.2 : Vector model of total angular momentum: spin and orbital coupling (spin-1/2 particles). (Public Domain; Maschen).

The result of these vector sums is specified in a code that is called a Russell-Saunders term symbol, and each term symbol identifies an energy level of the atom. Consequently, the energy levels also are called terms. A term symbol has the form $^{2s+1}L_J$ where the code letter that is used for the total orbital angular momentum quantum number L = 0, 1, 2, 3, 4, 5 is S, P, D, F, G, H, respectively. Note how this code matches that used for the atomic orbitals. The superscript $2S+1$ gives the spin multiplicity of the state, where S is the total spin angular momentum quantum number. The spin multiplicity is the number of spin states associated with a given electronic state. In order not to confuse the code letter S for the orbital angular momentum with the spin quantum number S, you must examine the context in which it is used carefully. In the term symbol, the subscript J gives the total angular momentum quantum number. Because of spin-orbit coupling, only J and M_j are valid quantum numbers, but because the spin-orbit coupling is weak L, M_l, S, and m_s still serve to identify and characterize the states for the lighter elements.

For example, the ground state, i.e. the lowest energy state, of the hydrogen atom corresponds to the electron configuration in which the electron occupies the 1s spatial orbital and can have either spin α or spin β. The term symbol for the ground state is $^2S_{1/2}$, which is read as “doublet S 1/2”. The spin quantum number is 1/2 so the superscript 2S+1 = 2, which gives the spin multiplicity of the state, i.e. the number of spin states equals 2 corresponding to α and β. The S in the term symbol indicates that the total orbital angular momentum quantum number is 0 (For the ground state of hydrogen, there is only one electron and it is in an s-orbital with $l = 0$). The subscript ½ refers to the total angular momentum quantum number. The total angular momentum is the sum of the spin and orbital angular momenta for the electrons in an atom. In this case, the total angular momentum quantum number is just the spin angular momentum quantum number, ½, since the orbital angular momentum is zero. The ground state has a degeneracy of two because the total angular momentum can have a z-axis projection of $+\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$, corresponding to m_J = +1/2 or -1/2 resulting from the two electron spin states α and β. We also can say, equivalently, that the ground state term or energy level is two-fold degenerate.

? Exercise 8.9.1

Write the term symbol for a state that has 0 for both the spin and orbital angular momentum quantum numbers.

? Exercise 8.9.2

Write the term symbol for a state that has 0 for the spin and 1 for the orbital angular momentum quantum numbers

---

8.9: The Allowed Values of J - the Total Angular Momentum Quantum Number is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.9: The Allowed Values of J - the Total Angular Momentum Quantum Number** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

8.10: Hund's Rules Determine the Term Symbols of the Ground Electronic States

Learning Objectives

- Define Hund three rules
- Use Hunds three rules to predict the lowest energy configuration and term symbols for multi-electron systems

The Aufbau section discussed how that electrons fill the lowest energy orbitals first, and then move up to higher energy orbitals only after the lower energy orbitals are full. However, there a problem with this rule. Certainly, 1s orbitals should be filled before 2s orbitals, because the 1s orbitals have a lower value of n, and thus a lower energy. What about the three different 2p orbitals? In what order should they be filled? The answer to this question involves Hund's rule, which make a lot more sense in the context of generated term symbols that are used to combine the various L and S values represent vector additions of possible microstates.

Hund's Rules

1. State with the largest value of S is most stable and stability decreases with decreasing S.
2. For states with same values of S, the state with the largest value of L is the most stable.
3. If states have same values of L and S then, for a subshell that is less than half filled, state with smallest J is most stable; for subshells that are more than half filled, state with largest value of J is most stable.

Example 8.10.1

Rank these terms associated with an electronic configuration of an atom based on energy (via Hund's rules):

$$^3D,\ ^3P,\ ^3S,\ ^1D,\ ^1P,\ ^1S$$

Hund's First Rule (Maximize Spin Multiplicity)

According to the first rule, electrons always enter an empty orbital before they pair up. Electrons are negatively charged and, as a result, they repel each other. Electrons tend to minimize repulsion by occupying their own orbitals, rather than sharing an orbital with another electron. Furthermore, quantum-mechanical calculations have shown that the electrons in singly occupied orbitals are less effectively screened or shielded from the nucleus.

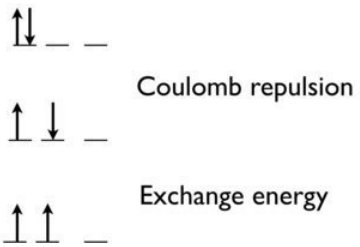


Figure 8.10.1 : Origin of Hund's First Rule. Hund's first rule is that for a given electron configuration, a state in which the spins are unpaired has lower total energy than one in which the spins are paired.

There's a Coulomb repulsion between two electrons to put them in the same orbital (a spin pairing energy often discussed in Crystal Field Theory). However, there's also a quantum mechanical effect. The *exchange energy* (which is favorable) increases with the number of possible exchanges between electrons with the same spin and energy. In transitioning from the top state to the middle state of Figure 8.10.1 , we remove the Coulomb repulsion between electrons in the same orbital. Moreover, In transitioning from the middle state to the bottom state (most stable state predicted by Hund's first rule), we *gain* the *exchange energy*, because these two electrons are indistinguishable.

Hund's Second Rule (Maximize Orbital Angular Multiplicity)

What matters is the total (scalar) angular momentum, not the direction. The negative and positive signs refer only to the direction of the angular momentum, not the magnitude. The direction is furthermore arbitrary (except in, say, a magnetic or electric field). So is the spin direction, incidentally. By convention we usually draw the first electron in each orbital as "up" (positive spin). However we could just as easily draw it "down". It makes no difference - in the absence of an external EM field, the energy is the same, if

only because molecules/atoms are rotating with respect to the lab frame anyway. "Up" and "down", in other words, is artificial. What matters is the *relative* momentum vectors of the various electrons in the system, and hence their sum total.

Hund's Third Rule (Minimize less than half filled or maximize greater than half filled shells)

A long time ago someone offering a reasonably simple explanation related to the fact that when the shell is more than half full, it's easier to visualize the system as an interaction between the spin and orbital momenta of holes rather than electrons, in which case the energetic stabilization term is reversed in sign. This would be because the spin angular momentum of a single hole would be opposite in sign compared to the spin angular momentum of a single electron. Taking as an example - the three p-orbitals. A situation with 1 electron and 5 electrons are functionally similar, except that one has a single electron and one has a single hole. All things being equal, the total spin angular momentum of the 1 electron system would be opposite in sign to whatever the total spin angular momentum of the 5 electron system is. So the expectations for Hund's rules would be switched. You can kind of see this if you draw out all the microstates of the 1-electron and 5-electron configurations: the everything is pretty much changed in sign in the latter case.

Example 8.10.2

What terms and levels can arise from an atom with the ground-state configuration of $1s^2 2s^2 2p^6 3s^2 3p^6 4s^2 4p^1 3d^1$? Which is the most stable (lowest in energy) state?

Solution

Possible states include:

$$^1F_3,\ ^1D_2,\ ^1P_1,\ ^3F_4,\ ^3F_3,\ ^3F_2,\ ^3D_3,\ ^3D_2,\ ^3D_1,\ ^3P_2,\ ^3P_1,\ ^3P_0.$$

There are two unpaired electrons in this system from the electron configuration.

- Rule 1 predicts that the ground state will be a triplet with $S = 1$ so $2S + 1 = 3$. So the ground state is from this more narrowed list: $^3F_4,\ ^3F_3,\ ^3F_2,\ ^3D_3,\ ^3D_2,\ ^3D_1,\ ^3P_2,\ ^3P_1,\ ^3P_0$.
- Rule 2 predicts a F state since that is the highest multiplicity with $L = 3$: So the ground state is from this more narrowed list: $^3F_4,\ ^3F_3,\ ^3F_2$
- Rule 3 predicts the lowest J term since the d shell is less than half full. That is the $J = 2$ state.

Therefore for this system, the atom will have a ground-state structure of

$$^3F_2$$

Example 8.10.3 : Titanium cation

The ground configuration of a Ti^{2+} ion is $[Ar]3d^2$. What is the term of lowest energy state?

Solution

- Rule 1: two unpaired electrons ⇒ highest S = 1 ⇒ 2S + 1 = 3
- Rule 2: two in d in parallel spin ⇒ highest L = 1 + 2 = 3 ⇒ 3F
- Rule 3:.L = 3, S = 1 ⇒ J = 4, 3, 2; less than half-filled ⇒ $^3F_2$

Exercise 8.10.3

What is the term of lowest energy state for the following atoms and ions.

- C: $[He]2s^2 2p^2$
- N: $[He]2s^2 2p^3$
- O: $[He]2s^2 2p^4$
- Cr^{3+}: $[Ar]3d^3$
- Mn^{3+}: $[Ar]3d^4$
- Fe^{3+}: $[Ar]3d^5$

8.10: Hund's Rules Determine the Term Symbols of the Ground Electronic States is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.10: Hund's Rules Determine the Term Symbols of the Ground Electronic States** by Delmar Larsen is licensed CC BY-NC-SA 4.0.

8.11: Using Atomic Term Symbols to Describe Atomic Spectra

Learning Objectives

- Demonstrate how spin-orbit coupling is experimentally observed in atomic spectra
- Use atomic terms symbols to ascribe transitions to specific angular momenta states described by atomic term symbols

Around 1930, several spectroscopists using high resolution instruments found that lines in the hydrogen atom spectrum actually are not single lines but they are multiplets as shown for an isotopic mixture of hydrogen,(H_α^1) and deuterium, (H_α^2), in Figure 8.11.1 . A multiplet consists of two or more closely spaced lines. Two lines together form a doublet, three a triplet, etc. Multiplets also are called fine structure. The term fine structure means the lines are spaced close together, i.e. finely spaced. Such fine structure also was found in spectra of one-electron ions such as $\mathbf{He^+}$.

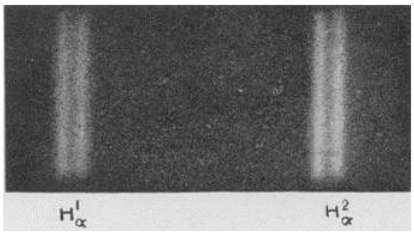


Figure 8.11.1 : Photograph of the first line in the Balmer series for atomic hydrogen and deuterium. These lines are identified as H_α^1 (hydrogen) and H_α^2 (deuterium), respectively. From H.E. White, Introduction to Atomic Spectra (McGraw-Hill, New York, 1934) p. 132 and G.N. Lewis and F.H. Spedding, Phys. Rev. 43, 964 (1933). The doublet splitting for H_α^1 at 656.279 nm was measured to be 0.326 cm^{-1}.

You should recall that the H_α^1 line in the Balmer series at 656.279 nm was understood as resulting from a single transition of an electron from the n = 3 energy level to the n = 2 level. The observation of fine structure revealed that an orbital energy level diagram does not completely describe the energy levels of atoms. This fine structure also provided key evidence at the time for the existence of electron spin, which was used not only to give a qualitative explanation for the multiplets but also to furnish highly accurate calculations of the multiplet splittings.

Spin-Orbit Coupling

Specifying the orbital configuration of an atom does not uniquely identify the electronic state of the atom because the orbital angular momentum, the spin angular momentum, and the total angular momentum are not precisely specified. For example in the hydrogen $2p^1$ configuration, the electron can be in any of the three p-orbitals, m_l = +1, 0, and –1, and have spins with m_s = +1/2 or –1/2. Thus, there are 3 times 2 different possibilities or states. Also, the orbital and spin angular momentum of the electrons combine in multiple ways to produce angular momentum vectors that are characteristic of the entire atom not just individual electrons, and these different combinations can have different energies. This coupling of orbital and spin angular momentum occurs because both the electron spin and orbital motion produce magnetic dipole moments. As we have seen previously, the relationship between the angular momentum and the magnetic moment is given by the gyromagnetic ratio. These magnetic dipoles interact just like two tiny bar magnets attracting and repelling each other. This interaction is called **spin-orbit interaction**. The interaction energy is proportional to the scalar product of the magnetic dipole moments, which are proportional to the angular momentum vectors.

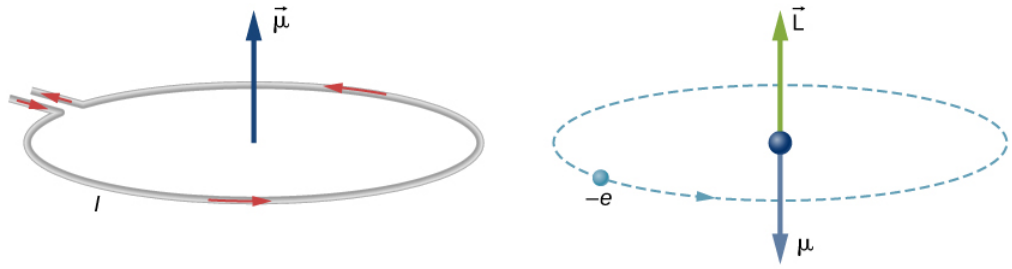
$$E_{s-o} \propto S \cdot L$$

with the following terms added to the Hamiltonian

$$\hat{H}_{s-o} \propto \hat{S} \cdot \hat{L}$$

where the constant of proportionality is called the spin-orbit coupling constant. The spin-orbit interaction couples the spin motion and orbital motion of all the electrons together. This coupling means that exact wavefunctions are not eigenfunctions of the spin and orbital angular momentum operators separately. Rather the total angular momentum $J = L + S$, the vector sum of the spin and orbital angular momentum, is required to be coupled for a completely accurate description of the system. Trying to describe the

coupled system in terms of spin and orbital angular momentum separately is analogous to trying to describe the positions of two coupled bar magnets independently. It cannot be done; their interaction must be taken into account (Figure 8.11.1).



(a) Current-carrying loop (b) Hydrogen atom

Figure 8.11.1 : An electron that orbits the nucleus induces a "current" and a magnetic field. Similarly, the rotating electron (really intrinsic spin) of the electron induces a different magnet. Both a magnetic (dipole) moments will interact with each other via spin-orbit coupling to generate fine structure atomic spectra.(CC-SA OpenStax).

Atomic Spectroscopy

Higher energy or excited orbital configurations also exist. The hydrogen atom can absorb energy, and the electron can be promoted to a higher energy orbital. The electronic states that result from these excited orbital configurations also are characterized or labeled by term symbols. The details of how to determine the term symbols for multi-electron atoms and for cases where both the orbital and spin angular momentum differ from zero are given elsewhere, along with rules for determining the relative energies of the terms.

We have found that the selection rules for promoting a single electron moving from one atomic orbital to another via the absorption or emission of light are

$$\Delta l = \pm 1$$
$$\Delta m_l = 0, \pm 1$$

These selection rules arise from the conservation fo angular momentum during a spectroscopic transition and the fact that a photon has a spin 1. Within the limits of L-S coupling, these rules can be expressed in terms of atomic term symbols resulting in the resulting Russell-Saunders selection rules:

$$\Delta S = 0 \tag{8.11.1}$$
$$\Delta L = 0, \pm 1 \tag{8.11.2}$$
$$\Delta J = 0, \pm 1, \tag{8.11.3}$$

but the $J = 0$ to $J = 0$ transition is forbidden

$$\Delta m_J = 0, \pm 1 \tag{8.11.4}$$

but the $m_J = 0$ to $m_J = 0$ transition is forbidden if $\Delta J = 0$.

These selection rules result from the general properties of angular momentum such as the conservation of angular momentum and commutation relations. The $\Delta L = 0$ option in Equation 8.11.2 does not violate the conservation of angular momentum discussed previously, since $\Delta l = \pm 1$ is still required. The orbital angular momentum of an electron must change upon absorption, but this does not necessarily affect the overall momentum of the state given by Equation 8.11.2.

The selection rules apply only to atoms that can be described with Russell-Saunders (LS) coupling. These rules fail as the atomic number increases because the S and L quantum numbers become "bad" quantum numbers; this occurs when the jj-coupling coupling approach is more applicable. For example, the transition between single ($S = 1/2$ and triplet $S = 1$ states (violation of selection rule in Equation 8.11.1) are allowed and experimentally observed, in heavy atoms.

✔ Example 8.11.1 : Sodium Atoms

An example of this fine structure is the emission of sodium atoms.



Figure above does not show this splitting).

616.07 nm	615.42 nm	589.00 nm	589.59 nm	568.82 nm	568.26 nm

How can these transitions be described in terms of transitions between microstates

Solution

We need to discussed states in terms of not only electron configurations, but in terms of microstates (i.e., term symbols) and the principal quantum number of the valence electron, n:

- The ground state has a $(Ne]ns^1$ configuration, which has only one microstate $^2S_{1/2}$
- The excited state with the valence electron in the p-orbitals has an electron configuration of $[Ne]np^1$, which has two microstates: $^2P_{3/2}$ and $^2P_{1/2}$
- The excited state with the valence electron in the p-orbitals has an electron configuration of $[Ne]nd^1$, which has two microstates of $^2D_{5/2}$ and $^2D_{3/2}$

Observed lines can be explained:

- $5S \rightarrow 3P$ gives two lines since the initial configuration has two microstates: 616.07, 615.42 nm
- $3P \rightarrow 3S$ gives two lines since the terminal configuration has two microstates: 589.00, 589.59 nm
- $4D \rightarrow 3P$ gives two lines since the terminal configuration has two microstates: 568.82, 568.26 nm

Splitting of the Sodium D Line

One notable atomic spectral line of sodium vapor is the so-called D-line, which may be observed directly as the sodium flame-test line and also the major light output of low -pressure sodium lamps (these produce pressure sodium lamps (these produce an unnatural yellow). The D-line is one of the classified Fraunhofer lines in Sodium vapor in the upper layers of lines. Sodium vapor in the upper layers of the sun creates a dark line in the emitted spectrum of electromagnetic radiation by absorbing visible light in a band of wavelengths around 589.5 nm. This wavelength corresponds to transitions in atomic sodium in which the valence-electron transitions from a 3s to 3p electronic state.

Closer examination of the visible spectrum of atomic sodium reveals that the D-line actually consists of two lines called the D_1 and D_2 lines at 589 6 nm and 589.0 nm, respectively. The splitting between these lines arises because of spin-orbit coupling. Na has one unpaired electron ($S = ½$). If we consider the $S \rightarrow P$ transition, then for the excited state, P, we have $L = 1$. Thus, $J = 3/2$ or $J = 1/2$.

Now we want to apply these ideas to understand why multiplet structure is found in the luminescence spectrum of hydrogen and single electron ions. As we have said, the H_α line in the Balmer series at 656.279 nm can be understood via a transition of an electron in a n = 3 atomic orbital to a n = 2 atomic orbital. When this spectral line was examined using high-resolution instruments, it was found actually to be a doublet, i.e. two lines separated by 0.326 cm^{-1}.

There are 9 degenerate orbitals associated with the n = 3 level, and 4 associated with the n = 2 level. Since an electron can be in any orbital with any one of two spins, we expect the total number of states to be twice the number of orbitals. The number of orbitals is given by n^2 so there should be 8 states associated with n = 2 and 18 states associated with n = 3. Using the ideas of vector addition of angular momentum, the terms that result from having an electron in any one of these orbitals are given in Table 8.11.1 .

Table 8.11.1 : H Atom Terms Originating from n = 1, 2, and 3

Orbital Configuration	Term Symbols	Degeneracy
$1s^1$	$^2S_{1/2}$	2
$2s^1$	$^2S_{1/2}$	2
$2p^1$	$^2P_{1/2}$, $^2P_{3/2}$	2, 4
$3s^1$	$^2S_{1/2}$	2
$3p^1$	$^2P_{1/2}$, $^2P_{3/2}$	2, 4
$3d^1$	$^2D_{3/2}$, $^2D_{5/2}$	4, 6

Table 8.11.1 shows that there are three terms associated with n = 2, and 5 terms associated with n = 3. In principle, each term can have a different energy. The degeneracy of each term is determined by the number of projections that the total angular momentum vector has on the z-axis. These projections depend on the m_J quantum number, which ranges from $+J$ to $-J$ in integer steps. J is the total angular momentum quantum number, which is given by the subscript in the term symbol. This relationship between m_J and J (m_J varies from $+J$ to $-J$ in integer steps) is true for any angular momentum vector.

? Exercise 8.11.3

Confirm that the nine term symbols in Table 8.11.1 are correct.

? Exercise 8.11.4

Confirm that the values for the degeneracy in Table 8.11.1 are correct and that the total number of states add up to 8 for n = 2 and 18 for n = 3.

The energies of the terms depend upon spin-orbit coupling and relativistic corrections that need to be included in the Hamiltonian operator in order to provide a more complete description of the hydrogen atom. As a consequence of these effects, all terms with the same n and (J\) quantum numbers have the **same energy,** while terms with different values for n or J have different energies. The theoretical term splittings as given by H.E. White, Introduction to Atomic Spectra (McGraw-Hill, New York, 1934) pp. 132-137 and are shown in Figure 8.11.2 .

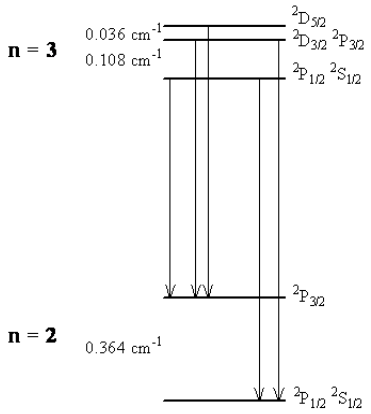


Figure 8.11.2 : Energy level diagram for the electronic states of hydrogen corresponding to an electron with the principal quantum number n = 3 and n = 2. While the relative energies of the terms are drawn to scale, the energy of n = 3 relative to n = 2 is not. Transitions to the $2P_{3/2}$ and ($2P_{1/2}$ and $2S_{1/2}$) levels correspond to the two lines in the H_α^1 band, which have a measured separation of 0.326 cm^{-1}.

Figure 8.11.2 shows 5 allowed transitions for the electron in the states associated with n = 3 to the states associated with n = 2. Of these five, two are most intense and are responsible for the doublet structure. These two transitions are indicated by the wide black lines at the bottom of the figure to correspond to the lines observed in the photographic spectrum shown in Figure 8.11.2 . The

other transitions contribute to the width of these lines or are not observed. The theoretical value for the doublet splitting is 0.328 cm^{-1}, which is in excellent agreement with the measured value of 0.326 cm^{-1}. The value of 0.328 cm^{-1} is obtained by taking the difference, 0.364 – 0.036 cm^{-1},in the term splittings.

As we have just seen, the electronic states, as identified by the term symbols, are essential in understanding the spectra and energy level structure of atoms, but it also is important to associate the term symbols and states with the orbital electron configurations. The orbital configurations help us understand many of the general or coarse features of spectra and are necessary to produce a physical picture of how the electron density changes because of a spectroscopic transition.

? Exercise 8.11.5

Use the Russell-Saunders selection rules to determine which transitions contribute to the H_α line in the hydrogen spectrum.

8.11: Using Atomic Term Symbols to Describe Atomic Spectra is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **8.11: Using Atomic Term Symbols to Interpret Atomic Spectra** by David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski is licensed CC BY-NC-SA 4.0.

8.E: Multielectron Atoms (Exercises)

Solutions to select questions can be found online.

8.4

Prove the speed of electron in first Bohr orbit is $e^2/4\pi\epsilon_0\hbar = 2.188 \times 10^6 m \bullet s^{-1}$. The speed is in atomic units.

8.15

Explain why:

$$E = \frac{\int \Phi_2^*(1,2)\hat{H}\Phi_2(1,2)dr_1dr_2d\sigma_1d\sigma_2}{\int \Phi_2^*(1,2)\Phi_2(1,2)dr_1dr_2d\sigma_1d\sigma_2}$$

can be rewritten as:

$$E = \frac{\int \Phi_2^*(1,2)\hat{H}\Phi_2(1,2)dr_1dr_2}{\int \Phi_2^*(1,2)\Phi_2(1,2)dr_1dr_2}$$

8.16

Why must you distinguish the two electrons in separated hydrogen atoms?

8.17

In both the Hartree-Fock approximation and hydrogen atomic wavefunction, why is the angular dependence the same?

8.20

Given the two electron determinate wavefunction below, determine if the spin component of the system is symmetric, anti-symmetric, or neither.

$$\begin{vmatrix} 1s\alpha(1) & 1s\beta(1) \\ 1s\alpha(2) & 1s\alpha(2) \end{vmatrix}$$

8.23

Given $\hat{S}_z\alpha = \frac{\hbar}{2}\alpha$ and $\hat{S}_z\alpha = \frac{-\hbar}{2}\beta$, show that Ψ_{200} is an eigenfunction of $S_{z,total} = \hat{S}_{z\alpha} + \hat{S}_{zB}$.

8.24

For the wavefunction:

$$\psi = \begin{vmatrix} \psi_A(1) & \psi_A(2) \\ \psi_B(1) & \psi_B(2) \end{vmatrix}$$

discuss the effect on the wavefunction of (a) swapping rows of the matrix and (b) swapping columns of the matrix.

8.27

What are the term symbols for carbon and oxygen atoms in the ground state?

8.28

Show that the number of sets of magnetic quantum number (m_l) and spin quantum number (m_s) associated with any term symbol is equal to $(2L+1)(2S+1)$. Apply this result to the np^2 case and show that symbol ^1S, ^3P and ^1D account for all the possible sets of magnetic quantum numbers and spin quantum numbers

8.29

Calculate all possible numbers of term symbol for an np^1 electron configuration.

8.30

Determine the ground state term symbol for the electron configuration of the Halogens.

8.33

^2P, ^2D, and ^4S are the term symbols for an atom with the np^3 electron configuration. Using the term symbols for the np^3 electron configuration, calculate the J values associated with each of the term symbol. Then find out which term symbol represent the ground state.

8.34

What are the ground state electron configuration and term symbol for Calcium?

8.34

Find the ground state term symbol for Ca.

8.36

Write the electron configuration for vanadium and use this information to find the ground-state term symbol for V.

8.36

What is the ground-state term symbol for Ne.

8.37

Looking at the $1s2p$ electron configuration for He. Solve for the term symbols (states) of the Helium configuration and the degeneracies. If electron spin orbit coupling is included, what effect will this have?

8.E: Multielectron Atoms (Exercises) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

9: The Chemical Bond: Diatomic Molecules

9.1: The Born-Oppenheimer Approximation Simplifies the Molecular Schrödinger Equation
9.2: The H_2^+ Prototypical Species
9.2B: Solving the H_2^+ System Exactly (Optional)
9.3: The Overlap Integral
9.3B: Evaluating the Overlap Integral (Optional)
9.4: Chemical Bond Stability
9.5: Bonding and Antibonding Orbitals
9.6: A Simple Molecular-Orbital Treatment of H_2 Places Both Electrons in a Bonding Orbital
9.7: Molecular Orbitals Can Be Ordered According to Their Energies
9.8: Molecular-Orbital Theory Does not Predict a Stable Diatomic Helium Molecule
9.9: Electrons Populate Molecular Orbitals According to the Pauli Exclusion Principle
9.10: Molecular Orbital Theory Predicts that Molecular Oxygen is Paramagnetic
9.E: The Chemical Bond: Diatomic Molecules (Exercises)

9: The Chemical Bond: Diatomic Molecules is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.1: The Born-Oppenheimer Approximation Simplifies the Molecular Schrödinger Equation

Learning Objectives

- Understand the need to introduce an approximation like the Born-Oppenheimer approximation to solve multi-electron systems
- Understand the basis of parameterization involved in using the Born-Oppenheimer approximation

Using quantum mechanics to predict the chemical bonding patterns, optimal geometries, and physical and chemical properties of molecules is a large and active field of research known as *molecular quantum mechanics* or more commonly as *quantum chemistry*. The density functional theory referred to in the previous lecture, for which the chemistry Nobel prize was given in 1998, has had a tremendous impact in quantum chemistry, with some of the papers in this subject having acquired some 10,000 citations each since their publication. In fact, the 1998 chemistry Nobel prize was shared between Walter Kohn, one of the inventors of density functional theory and John Pople, the developer of a commonly used quantum chemistry software package.

Quantum chemistry calculations allow the geometries of molecules to be computed as well as a wide range of properties. Quantum chemistry can also be used in a novel way, in which the electrons are treated using quantum mechanics but the nuclei are treated as classical particles. We use quantum mechanics to calculate the internuclear forces but then use these forces in Newton's Second Law to study the motion of the nuclei during chemical reactions. This gives us a microscopic window into the specific motions, the complex dance, executed by the nuclei during a simple or complex chemical process.

The methods of quantum chemistry have become very sophisticated, and there are various software packages that can be downloaded for carrying out the calculations of quantum chemistry. It should be noted that these packages use a series of approximations to solve the Schrödinger equation because for all but the simplest of molecules, exact solutions are not available. We will discuss some of these methods, but first we need to introduce some of the underlying theory.

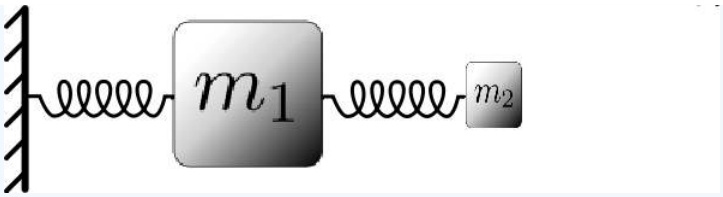
The Born-Oppenheimer Approximation

The Born-Oppenheimer approximation is one of the basic concepts underlying the description of the quantum states of molecules. This approximation makes it possible to separate the motion of the nuclei and the motion of the electrons. This is not a new idea for us. We already made use of this approximation in the particle-in-a-box model when we explained the electronic absorption spectra of cyanine dyes without considering the motion of the nuclei. Then we discussed the translational, rotational and vibrational motion of the nuclei without including the motion of the electrons. In this chapter we will examine more closely the significance and consequences of this important approximation. Note, in this discussion nuclear refers to the atomic nuclei as parts of molecules not to the internal structure of the nucleus.

The Born-Oppenheimer approximation neglects the motion of the atomic nuclei when describing the electrons in a molecule. The physical basis for the Born-Oppenheimer approximation is the fact that the mass of an atomic nucleus in a molecule is much larger than the mass of an electron (more than 1000 times). Because of this difference, the nuclei move much more slowly than the electrons. In addition, due to their opposite charges, there is a mutual attractive force of Ze^2/r^2 acting on an atomic nucleus and an electron. This force causes both particles to be accelerated. Since the magnitude of the acceleration is inversely proportional to the mass, a = F/m, the acceleration of the electrons is large and the acceleration of the atomic nuclei is small; the difference is a factor of more than 2000. Consequently, the electrons are moving and responding to forces very quickly, and the nuclei are not. You can imagine running a 100-yard dash against someone whose acceleration is a 2000 times greater than yours. That person could literally run circles around you.

Example 9.1.1 : Coupled Oscillators with Dissimilar Masses

If two particles interact in some way, and one is much heavier than the other, the light particle will move essentially as a "slave" of the heavy particle. That is, it will simply follow the heavy particle wherever it goes, *and*, it will move rapidly in response to the heavy particle motion. As an illustration of this phenomenon, consider the simple mechanical system pictured below:



(CC BY-SA 3.0 Unported; Jim.belk via Wikipedia).

Considering this as a classical system, we expect that the motion will be dominated by the large heavy particle (m_1), which is attached to a fixed wall by a spring. The small, light particle (m_2, which is attached to the heavy particle by a spring will simply follow the heavy particle and execute rapid oscillations around it.

So a good approximation is to describe the electronic states of a molecule by thinking that the nuclei are not moving, i.e. that they are stationary. The nuclei, however, can be stationary at different positions so the electronic wavefunction can depend on the positions of the nuclei even though their motion is neglected.

Now we look at the mathematics to see what is done in solving the Schrödinger equation after making the Born-Oppenheimer approximation. For a diatomic molecule as an example, the Hamiltonian operator is grouped into three terms

$$\hat{H}(r,R) = \hat{T}_{nuc}(R) + \frac{e^2}{4\pi\epsilon_0}\frac{Z_A Z_B}{R} + \hat{H}_{elec}(r,R) \tag{9.1.1}$$

where

$$\hat{T}_{nuc}(R) = -\frac{\hbar^2}{2m_A}\nabla_A^2 - \frac{\hbar^2}{2m_B}\nabla_B^2 \tag{9.1.2}$$

and

$$\hat{H}_{elec}(\vec{r},\vec{R}) = \frac{-\hbar^2}{2m}\sum_i \nabla_i^2 + \frac{e^2}{4\pi\epsilon_0}\left(-\sum_i \frac{Z_A}{r_{Ai}} - \sum_i \frac{Z_B}{r_{Bi}} + \frac{1}{2}\sum_i \sum_{j\neq i} \frac{1}{r_{ij}}\right) \tag{9.1.3}$$

In Equation 9.1.1, the first term represents the kinetic energy of the nuclei, the second term represents the Coulomb repulsion of the two nuclei, and the third term represents the contribution to the energy from the electrons, which consists of their kinetic energy, mutual repulsion for each other, and attraction for the nuclei. $\vec{r}$ and $\vec{R}$ are vectors specifying the positions of all the electrons and all the nuclei, respectively.

? Exercise 9.1.1

Define all the symbols in Equations 9.1.1 through 9.1.3.

? Exercise 9.1.2

Explain why the factor of 1/2 appears in the last term in Equation 9.1.3.

The Born-Oppenheimer approximation says that the nuclear kinetic energy terms in the complete Hamiltonian, Equation 9.1.1, can be neglected in solving for the electronic wavefunctions and energies. Consequently, the electronic wavefunction $\varphi_e(r,R)$ is found as a solution to the electronic Schrödinger equation

$$\hat{H}_{elec}(r,R)\varphi_e(r,R) = E_e(R)\varphi_e(r,R) \tag{9.1.4}$$

Even though the nuclear kinetic energy terms are neglected, the Born-Oppenheimer approximation still takes into account the variation in the positions of the nuclei in determining the electronic energy and the resulting electronic wavefunction depends upon the nuclear positions, R. As a result of the Born-Oppenheimer approximation, the molecular wavefunction can be written as a product

$$\psi_{ne}(r,R) = X_{ne}(R)\varphi_e(r,R) \tag{9.1.5}$$

This product wavefunction is called the Born-Oppenheimer wavefunction. The function $X_{ne}(R)$ is the vibrational wavefunction, which is a function of the nuclear coordinates R and depends upon both the vibrational and electronic quantum numbers or states, n and e, respectively. The electronic function, $\varphi_e(r, R)$, is a function of both the nuclear and electronic coordinates, but only depends upon the electronic quantum number or electronic state, e. Translational and rotational motion is not included here. The translational and rotational wavefunctions simply multiply the vibrational and electronic functions in Equation 9.1.5 to give the complete molecular wavefunction when the translational and rotational motions are not coupled to the vibrational and electronic motion.

In the *Crude Born-Oppenheimer Approximation*, R is set equal to R_o, the equilibrium separation of the nuclei, and the electronic wavefunctions are taken to be the same for all positions of the nuclei (i.e., the nuclei never move). The electronic energy, $E_e(R)$, in Equation 9.1.4 combines with the repulsive Coulomb energy of the two nuclei, to form the potential energy function that controls the nuclear motion as shown in Figure 9.1.1 .

$$V_e(R) = E_e(R) + \frac{e^2}{4\pi\epsilon_0}\frac{Z_A Z_B}{R} \tag{9.1.6}$$

Consequently the Schrödinger equation for the vibrational motion is

$$(\hat{T}_{nuc}(R) + V(R))X_{ne}(R) = E_{ne}X_{ne}(R) \tag{9.1.7}$$

Potential Energy Curves and Surfaces

Previously, the potential energy was approximated as a harmonic potential or Morse potential depending on the displacement, R, of the nuclei from their equilibrium positions.

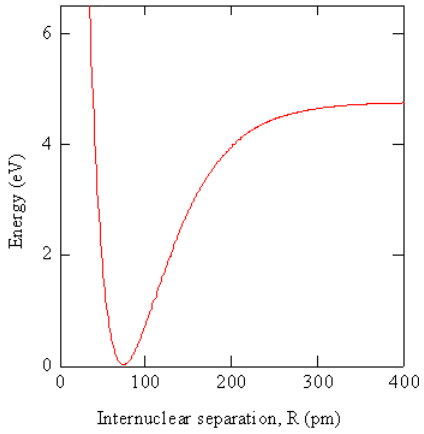


Figure 9.1.1 : The potential energy function for a diatomic molecule.

In practice the electronic Schrödinger equation is solved using approximations at particular values of R to obtain the wavefunctions $\varphi_e(r, R)$ and potential energies $V_e(R)$. The potential energies can be graphed as illustrated in Figure 9.1.1 .

The graph in Figure 9.1.1 is the energy of a diatomic molecule as a function of internuclear separation, which serves as the potential energy function for the nuclei. When R is very large there are two atoms that are weakly interacting. As R becomes smaller, the interaction becomes stronger, the energy becomes a large negative value, and we say a bond is formed between the atoms. At very small values of R, the internuclear repulsion is very large so the energy is large and positive. This energy function controls the motion of the nuclei. Previously, we approximated this function by a harmonic potential to obtain the description of vibrational motion in terms of the harmonic oscillator model. Other approximate functional forms could be used as well, e.g. the Morse potential. The equilibrium position of the nuclei is where this function is a minimum, i.e. at $R = R_0$. If we obtain the wavefunction at $R = R_0$ and use this function for all values of R, we have employed the *Crude Born-Oppenheimer* approximation.

? Exercise 9.1.3

Relate Equation 9.1.7 to the one previously used in our description of molecular vibrations in terms of the harmonic oscillator model.

While the potential energy function, $V_e(R)$, for a diatomic molecule is a 1-D curve (Figure 9.1.1), molecules with more than two atoms will have multi-dimensional **potential energy surfaces** with 3N-6 (or 3N-5 for linear molecule) dimensions for the number of internal degrees of freedom.

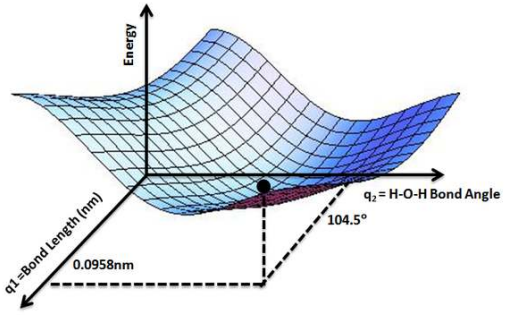


Figure 9.1.2 : The potential energy surface for a water molecule: Shows the energy minimum corresponding to optimized molecular structure for water- O-H bond length of 0.0958 nm and H-O-H bond angle of 104.5°. from Wikipeda (AimNature)

The potential energy surface concept can be used to theoretically explore properties of structures composed of atoms, for example, finding the minimum energy shape of a molecule or computing the rates of a chemical reaction. Qualitatively the reaction coordinate diagrams (one-dimensional slices through the potential energy surfaces) have numerous applications. Chemists use reaction coordinate diagrams as both an analytical and pedagogical aid for rationalizing and illustrating kinetic and thermodynamic events. The purpose of energy profiles and surfaces is to provide a qualitative representation of how potential energy varies with molecular motion for a given reaction or process.

? Exercise 9.1.4

Explain the difference between the Born-Oppenheimer approximation and the Crude Born-Oppenheimer approximation.

Summary

In this section we started with the Schrödinger equation for a diatomic molecule and separated it into two equations, an electronic Schrödinger equation and a nuclear Schrödinger equation. In order to make the separation, we had to make an approximation. We had to neglect the effect of the nuclear kinetic energy on the electrons. The fact that this assumption works can be traced to the fact that the nuclear masses are much larger than the electron mass. We then used the solution of the electronic Schrödinger equation to provide the potential energy function for the nuclear motion. The solution to the nuclear Schrödinger equation provides the vibrational wavefunctions and energies.

Contributors and Attributions

David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")

9.1: The Born-Oppenheimer Approximation Simplifies the Molecular Schrödinger Equation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **9.1: The Born-Oppenheimer Approximation Simplifies the Schrödinger Equation for Molecules** is licensed CC BY-NC-SA 4.0.

9.2: The H_2^+ Prototypical Species

Molecular orbital theory is a conceptual extension of the orbital model, which was so successfully applied to atomic structure. As was once playfully remarked, "a molecule is nothing more than an atom with more nuclei." This may be overly simplistic, but we do attempt, as far as possible, to exploit analogies with atomic structure. Our understanding of atomic orbitals began with the exact solutions of a prototype problem – the hydrogen atom. We will begin our study of homonuclear diatomic molecules beginning with another exactly solvable prototype, the hydrogen molecule-ion $\mathbf{H_2^+}$.

The Hydrogen Molecular Ion

The simplest conceivable molecule would be made of two protons and one electron, namely $\mathbf{H_2^+}$. This species actually has a transient existence in electrical discharges through hydrogen gas and has been detected by mass spectrometry and it also has been detected in outer space. The Schrödinger equation for $\mathbf{H_2^+}$ can be solved exactly within the Born-Oppenheimer approximation (i.e., fixed nuclei). This ion consists of two protons held together by the electrostatic force of a single electron. Clearly the two protons, two positive charges, repeal each other. The protons must be held together by an attractive Coulomb force that opposes the repulsive Coulomb force. A negative charge density between the two protons would produce the required counter-acting Coulomb force needed to pull the protons together. So intuitively, to create a chemical bond between two protons or two positively charged nuclei, a high density of negative charge between them is needed. We expect the molecular orbitals that we find to reflect this intuitive notion.

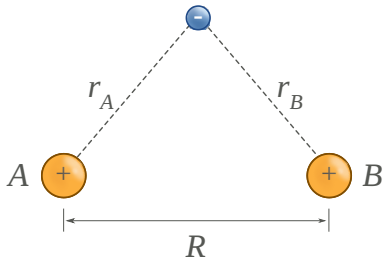


Figure 9.2.1 : Hydrogen molecular ion $\mathbf{H_2^+}$ with fixed nuclei $\boldsymbol{A}$ and $\boldsymbol{B}$, internuclear distance $\boldsymbol{R}$. (CC BY-NC; Ümit Kaya via LibreTexts)

The electronic Hamiltonian for $\mathbf{H_2^+}$ is

$$\hat{H}_{elec}(r,R) = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0 r_A} - \frac{e^2}{4\pi\epsilon_0 r_B} + \frac{e^2}{4\pi\epsilon_0 R} \tag{9.2.1}$$

where r_A and r_B are the distances electron from the $\boldsymbol{A}$ and $\boldsymbol{B}$ hydrogen nuclei, respectively and $\boldsymbol{R}$ is the distance between the two protons.

Although the Schrödinger equation for $\mathbf{H_2^+}$ can be solved exactly (albeit within the Born-Oppenheimer approximation where the nuclei are fixed) because there is only one electron, we will develop approximate solutions in a manner applicable to other diatomic molecules that have more than one electron.

Linear Combination of Atomic Orbitals

For the case where the protons in H_2^+ are infinitely far apart, we have a hydrogen atom and an isolated proton when the electron is near one proton or the other. The electronic wavefunction would just be $1s_A(r)$ or $1s_B(r)$ depending upon which proton, labeled $\boldsymbol{A}$ or $\boldsymbol{B}$, the electron is near. Here $1s_A$ denotes a 1s hydrogen atomic orbital with proton A serving as the origin of the spherical polar coordinate system in which the position r of the electron is specified. Similarly $1s_B$ has proton B as the origin. A useful approximation for the molecular orbital when the protons are close together therefore is a linear combination of the two atomic orbitals. The general method of using

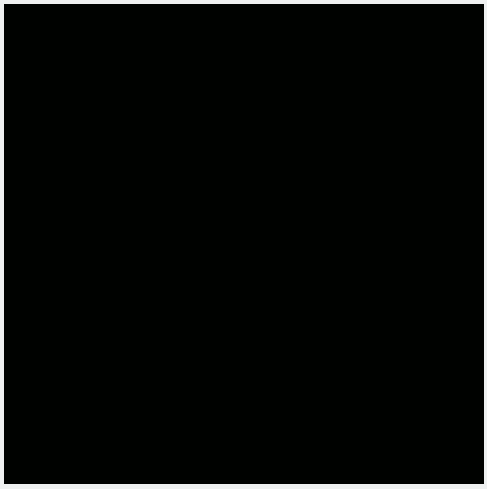
$$\psi(r) = C_A 1s_A(r) + C_B 1s_B(r) \tag{9.2.2}$$

i.e. of finding molecular orbitals as linear combinations of atomic orbitals is called the *Linear Combination of Atomic Orbitals - Molecular Orbital (LCAO-MO) Method*. In this case we have two basis functions in our basis set, the hydrogenic atomic orbitals $1s_A$ and $1s_B$.

The Linear Combination of Atomic Orbitals (LCAO) Approximation and Interference

The LCAO approximation is an example of the linear variational method discussed previously with the true molecular orbital wavefunction approximated as an expansion of a basis set of atomic orbitals on each atom of the molecule with variable coefficients that can be optimized (e.g., via the secular equations). As discussed previously, the number of wavefunctions (solutions) extracted from solving the secular determinant is equal the number of elements in the expansion. So for the expansion in Equation 9.2.2 with two atomic orbitals contributing result in two molecule orbitals.

This method yields a approximate picture of the molecular orbitals in a molecules. The figure below shows two atoms approaching along the axis of one of their $2p$ states. In the top row, the two lobes facing one another have the same sign; in the bottom row they have opposite sign. These are two different linear combinations of the same two atomic states, on different atoms and with difference phases (i.e., signs of C_A vs. C_B in the expansion). In the first example, the electron density increases between the nuclei and in the second example, a very steep-sided node between the two nuclei causes all the probability density to face away from the atom opposite.



For H_2^+, the simplest molecule, the starting function is given by Equation 9.2.2. We must determine the values for the coefficients, C_A and C_B. We could use the variational method to find a value for these coefficients, but for the case of H_2^+ evaluating these coefficients is easy. Since the two protons are identical, the probability that the electron is near A must equal the probability that the electron is near B. These probabilities are given by $|C_A|^2$ and $|C_B|^2$, respectively. Consider two possibilities that satisfy the condition $|C_A|^2 = |C_B|^2$; namely, $C_A = C_B = C_+$ and $C_A = -C_B = C_-$. These two cases produce two molecular orbitals:

$$\psi_+ = C_+(1s_A + 1s_B) \tag{9.2.3}$$

$$\psi_- = C_-(1s_A - 1s_B) \tag{9.2.4}$$

The probability density for finding the electron at any point in space is given by $|\psi^2|$ and the electronic charge density is just $|e\psi^2|$. The important difference between ψ_+ and ψ_- is that the charge density for ψ_+ is enhanced (Figure 9.2.2 (bottom) between the two protons, whereas it is diminished for ψ_- as shown in Figures 9.2.2 (top). ψ_- has a node in the middle while ψ_+ corresponds to our intuitive sense of what a chemical bond must be like. The electronic charge density is enhanced in the region between the two protons.

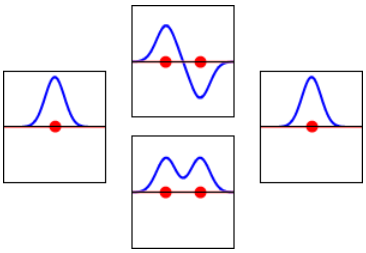


Figure 9.2.2 : Electron wavefunctions for the 1s orbital of a lone hydrogen atom (left and right) and the corresponding bonding (bottom) and antibonding (top) molecular orbitals of the H_2^+ ion. The real part of the wavefunction is the blue curve, and the imaginary part is the red curve. The red dots mark the locations of the nuclei. The electron wavefunction oscillates according to the Schrödinger wave equation, and orbitals are its standing waves. The standing wave frequency is proportional to the orbital's kinetic energy. (This plot is a one-dimensional slice through the three-dimensional .)

So ψ_+ is called a **bonding molecular orbital**. If the electron were described by ψ_-, the low charge density between the two protons would not balance the Coulomb repulsion of the protons, so ψ_- is called an **antibonding molecular orbital**.

Contributors and Attributions

- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")
- Rudolf Winter (Aberystwyth University)

9.2: The H_2^+ Prototypical Species is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **9.2: The H_2^+ Prototypical Species** is licensed CC BY-NC-SA 4.0.

9.2B: Solving the H_2^+ System Exactly (Optional)

The hydrogen molecule ion H_2^+ is the *only* molecule for which we can solve the electronic Schrödinger equation exactly (and only within the Born Oppenheimer Approximation). Note that it has just one electron! In fact, there are no multi-electron molecules we can solve exactly. Thus from H_2 on up to more complicated molecules, we only have approximate solutions for the allowed electronic energies and wavefunctions. Before we discuss these, however, let us examine the exact solutions for H_2^+ starting with a brief outline of how the exact solution is carried out.

Figure 9.2B.1 shows the geometry of the H_2^+ molecule ion and the coordinate system we will use.

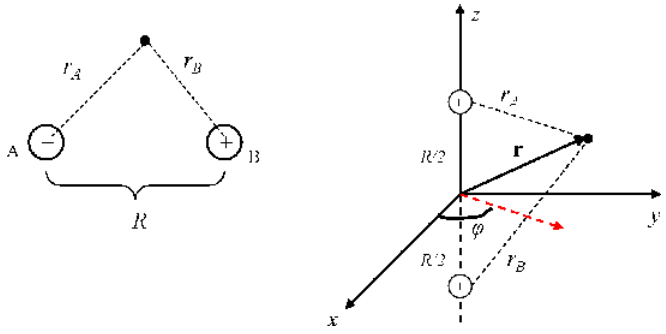


***Figure* 9.2B.1**

The two protons are labeled A and B, and the distances from each proton to the one electron are r_A and r_B, respectively. Let R be the distance between the two protons (this is the only nuclear degree of freedom that is important, and the electronic wavefunction will depend parametrically only on R). The coordinate system is chosen so that the protons lie on the z-axis one at a distance $R/2$ above the xy plane and one a distance $-R/2$ below the xy plane. The classical energy of the electron is

$$\frac{p^2}{2m_e} - \frac{e^2}{4\pi\epsilon_0}\left(\frac{1}{r_A} + \frac{1}{r_B}\right) = E \tag{9.2B.1}$$

The nuclear-nuclear term $V_{nn}(R) = e^2/4\pi\epsilon_0 R$ is a constant, and we can define the potential energy relative to this quantity.

The energy is not a simple function of energies for the x, y, and z directions, so we try another coordinate system to see if we can simplify the problem. In fact, this problem has a natural cylindrical symmetry (analogous to the spherical symmetry of the hydrogen atom) about the z-axis. Thus, we try cylindrical coordinates. In cylindrical coordinates the distance of the electron from the z-axis is denoted ρ, the angle ϕ is the azimuthal angle, as in spherical coordinates, and the last coordinate is just the Cartesian z coordinate. Thus,

- $x = \rho\cos\phi$
- $y = \rho\sin\phi$
- $z = z$

Using right triangles, the distance r_A and r_B can be shown to be

$$r_A = \sqrt{\rho^2 + (R/2 - z)^2} \quad r_B = \sqrt{\rho^2 + (R/2 + z)^2} \tag{9.2B.2}$$

The classical energy becomes

$$\frac{p_\rho^2}{2m_e} + \frac{p_\phi^2}{2m_e\rho^2} + \frac{p_z^2}{2m_e} - \frac{e^2}{4\pi\epsilon_0}\left(\frac{1}{\sqrt{\rho^2 + (R/2 - z)^2}} + \frac{1}{\sqrt{\rho^2 + (R/2 + z)^2}}\right) = E \tag{9.2B.3}$$

First, we note that the potential energy does not depend on ϕ, and the classical energy can be written as a sum $E = \varepsilon_{\rho,z} + \varepsilon_\phi$ of a ρ and z dependent term and an angular term. Moreover, angular momentum is conserved as it is in the hydrogen atom. However, in this case, only *one* component of the angular momentum is actually conserved, and this is the z-component, since the motion is completely symmetric about the z-axis. Thus, we have only certain allowed values of L_z, which are $m\bar{h}$, as in the hydrogen atom, where $m = 0, \pm1, \pm2, \ldots$. The electronic wavefunction (now dropping the "(elec)" label, since it is understood that we are discussing *only* the electronic wavefunction), can be written as a product

$$\psi(\rho,\phi,z) = G(\rho,z)y(\phi) \tag{9.2B.4}$$

and $y(\phi)$ is given by

$$y_m(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi} \tag{9.2B.5}$$

which satisfies the required boundary condition $y_m(0) = y_m(2\pi)$. Unfortunately, what is left in ρ and z is still not that simple. But if we make one more change of coordinates, the problem simplifies. We introduce two new coordinates μ and ν defined by

$$\mu = \frac{r_A + r_B}{R} \quad \nu = \frac{r_A - r_B}{R} \tag{9.2B.6}$$

Note that when $\nu = 0$, the electron is in the xy plane. Thus, ν is analogous to z in that it varies most as the electron moves along the z-axis. The presence or absence of a node in the xy plane will be an important indicator of wavefunctions that lead to a chemical bond in the molecule or not. The coordinate μ, on the other hand, is minimum when the electron is on the z-axis and grows as the distance of the electron from the z-axis increases. Thus, μ is analogous to ρ. The advantage of these coordinates is that the wavefunction turns out to be a simple product

$$\psi(\mu,\nu,\phi,R) = M(\mu)N(\nu)y(\phi) \tag{9.2B.7}$$

which greatly simplifies the problem and allows the exact solution.

The mathematical structure of the exact solutions is complex and nontransparent, so we will only look at these graphically, where we can gain considerably insight. First, we note that the quantum number m largely determines how the solutions appear. First, let us introduce the nomenclature for designating the orbitals (solutions of the Schrödinger equation, wavefunctions) of the system

- If $m = 0$, the orbitals are called σ orbitals, analogous to the s orbitals in hydrogen.
- If $m = 1$, the orbitals are called π orbitals, analogous to the p orbitals in hydrogen.
- If $m = 2$, the orbitals are called δ orbitals, analogous to the d orbitals in hydrogen.
- If $m = 3$, the orbitals are called ϕ orbitals, analogous to the f orbitals in hydrogen.

These orbitals are known as *molecular orbitals* because the describe the electronic wavefunctions of an entire molecule. There are four designators that we use to express each molecular orbital:

I. A greek letter, σ, π, δ, ϕ, ... depending on the quantum number m.
II. A subscript qualifier g or u depending on how an orbital ψ behaves with respect to a spatial reflection or *parity* operation $r \rightarrow -r$. If

$$\psi(-r) = \psi(r) \tag{9.2B.8}$$

then $\psi(r)$ is an even function of r, so we use the $"g"$ designator, where g stands for the German word *gerade*, meaning "even". If

$$\psi(-r) = -\psi(r) \tag{9.2B.9}$$

then $\psi(r)$ is an odd function of r, and we use the $"u"$ designator, where u stands for the German word *ungerade*, meaning "odd".
III. An integer n in front of the Greek letter to designate the energy level. This is analogous to the integer we use in atomic orbitals $(1s, 2s, 2p, ...)$.
IV. *An asterisk or no asterisk depending on the presence or absence of nodes between the nuclei. If there is significant amplitude between the nuclei, then the orbital favors a chemical bond, and the orbital is called a bonding orbital. If there is a node between the nuclei, the orbital does not favor bonding, and the orbital is called an antibonding orbital.*

So, the first few orbitals, in order of increasing energy are:

$$1\sigma_g, 1\sigma_u^*, 2\sigma_g, 2\sigma_u^*, 1\pi_u, 3\sigma_g, 1\pi_g^*, 3\sigma_u \tag{9.2B.10}$$

Contributors

- Mark Tuckerman (New York University)

9.2B: Solving the H_2^+ System Exactly (Optional) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.3: The Overlap Integral

For H_2^+, the simplest molecule, we must determine the values for the coefficients, C_A and C_B for the LCAO approximation for the molecular orbital as a linear combination of the two atomic orbitals

$$|\psi(r)\rangle = C_A 1s_A(r) + C_B 1s_B(r)$$

We could use the linear variational method to find a value for these coefficients, but for the case of H_2^+ evaluating these coefficients is easy. Since the two protons are identical, the probability that the electron is near A must equal the probability that the electron is near B. These probabilities are given by $|C_A|^2$ and $|C_B|^2$, respectively. Consider two possibilities that satisfy the condition

$$|C_A|^2 = |C_B|^2$$

so $C_A = C_B = C_+$ and $C_A = -C_B = C_-$. These two cases produce two molecular orbitals:

$$\underbrace{|\psi_+\rangle = C_+(1s_A + 1s_B)}_{\text{bonding molecular orbital}} \tag{9.3.1}$$

and

$$\underbrace{|\psi_-\rangle = C_-(1s_A - 1s_B)}_{\text{antibonding molecular orbital}} \tag{9.3.2}$$

The probability density for finding the electron at any point in space is given by $|\psi^2|$ and the electronic charge density is just $|e\psi^2|$. The important difference between ψ_+ and ψ_- is that the charge density for $|\psi_+\rangle$ is enhanced between the two protons, whereas it is diminished for $|\psi_-\rangle$ as shown in Figures 9.3.1 . ψ_- has a node in the middle while $|\psi_+\rangle$ corresponds to our intuitive sense of what a chemical bond must be like. The electronic charge density is enhanced in the region between the two protons. So $|\psi_+\rangle$ is called a *bonding molecular orbital*. If the electron were described by $|\psi_-\rangle$, the low charge density between the two protons would not balance the Coulomb repulsion of the protons, so $|\psi_-\rangle$ is called an *antibonding molecular orbital*.

Now we want to evaluate C_+ and C_- and then calculate the energy. The bonding and antibonding character of ψ_+ and ψ_- also should be reflected in the energy. If $|\psi_+\rangle$ indeed describes a bonding orbital, then the energy of this state should be less than that of a proton and hydrogen atom that are separated. The calculation of the energy will tell us whether this simple theory predicts H_2^+ to be stable or not and also how much energy is required to dissociate this molecule.

The constants C_+ and C_- are evaluated from the standard normalization condition:

$$\int \psi_\pm^* \psi_\pm d\tau = \langle \psi_\pm | \psi_\pm \rangle = 1 \tag{9.3.3}$$

$$\langle C_\pm(1s_A \pm 1s_B) | C_\pm(1s_A \pm 1s_B) \rangle = 1 \tag{9.3.4}$$

$$|C_\pm|^2[\underbrace{\langle 1s_A|1s_A\rangle}_{1} + \underbrace{\langle 1s_B|1s_B\rangle}_{1} \pm \underbrace{\langle 1s_B|1s_A\rangle}_{S} \pm \underbrace{\langle 1s_A|1s_B\rangle}_{S^*}] = 1 \tag{9.3.5}$$

Since the atomic orbitals are normalized, the first two integrals are just 1. The last two integrals are called *overlap integrals* and are symbolized by S and S^*, respectively, since one is the complex conjugate of the other.

The overlap integrals are telling us to take the value of $1s_B$ at a point multiply by the value of $1s_A$ at that point and sum (integrate) such a product over all of space (Figure 9.3.1). If the functions do not overlap, i.e. if one is zero when the other one is not and vice versa, these integrals then will be zero. It also is possible in general for such integrals to be zero even if the functions overlap because of the cancelation of positive and negative contributions.

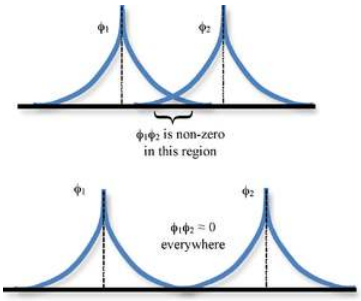


Figure 9.3.1 : The wavefunctions of atomic orbitals decrease exponentially with distance. Orbital overlap is non-zero when two atoms are close together, as illustrated for 1s orbitals in the upper figure. The lower figure shows orbitals that are too far away to interact. In this case both S is close to zero.

If the overlap integral is zero, for whatever reason, the functions are said to be *orthogonal*. Notice that the overlap integral ranges from 0 to 1 as the separation between the protons varies from $R = \infty$ to $R = 0$. Clearly when the protons are infinite distance apart, there is no overlap, and when $R = 0$ both functions are centered on one nucleus and $\langle 1s_A | 1s_B \rangle$ becomes identical to $\langle 1s_B | 1s_A \rangle$, which is normalized to 1, because $1s_A = 1s_B$.

With these considerations and using the fact that $1s$ wavefunctions are real so

$$\langle 1s_A | 1s_B \rangle = \langle 1s_B | 1s_A \rangle = S \tag{9.3.6}$$

Equation 9.3.5 becomes

$$|C_{\pm}|^2 (2 \pm 2S) = 1 \tag{9.3.7}$$

The solution to Equation 9.3.7 is given by

$$C_{\pm} = \frac{1}{\sqrt{2(1 \pm S)}} \tag{9.3.8}$$

Hence, the normalized molecular orbitals in Equations 9.3.1 and 9.3.1 are

$$|\psi_+\rangle = \frac{1}{\sqrt{2(1+S)}} (1s_A + 1s_B) \tag{9.3.9}$$

and

$$|\psi_-\rangle = \frac{1}{\sqrt{2(1-S)}} (1s_A - 1s_B) \tag{9.3.10}$$

The energies associated with these wavefunctions requires a bit more effort to calculate though as demonstrated in the following section.

For the overlap integral of two 1s orbitals from the hydrogen dimer discussed above is difficult to evaluate analytically and is explained here. The final answer is:

$$S(R) = \langle 1s_A | 1s_B \rangle \tag{9.3.11}$$

$$= e^{-R/a_o} \left(1 + \frac{R}{a_o} + \frac{R^2}{3a_o^2} \right) \tag{9.3.12}$$

The overlap integral for two 1s atomic orbitals of hydrogen is graphically displayed below

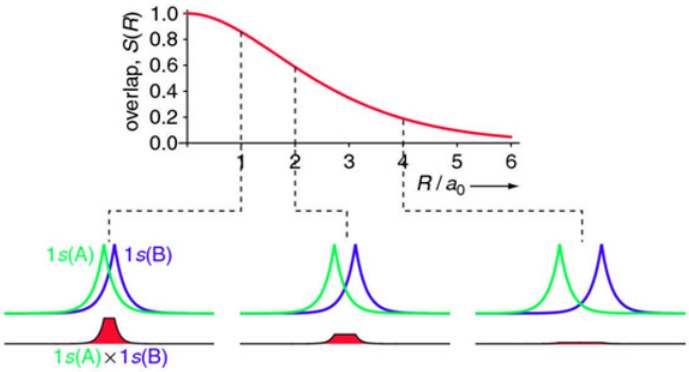


Figure 9.3.2 : The overlap integral at different proton separations.

Example 9.3.1

Calculate the difference in the electronic charge density (C/pm^3) at a point halfway between the two nuclei for $\mathbf{H_2^+}$ for an electron in the bonding molecular orbital compared to one in the antibonding molecular orbital.

Solution

The electronic charge density is calculated with the formula $|e\psi^2|$ where $e = 1.602 \times 10^{-19} C$

a: Bonding Molecular Orbital

First, squaring the bonding molecular orbital (Equation 9.3.9) gives us:

$$\begin{aligned}\psi_+^2 &= \frac{1}{2(1+S)}(1s_A + 1s_B)^2 \\ &= \frac{1}{2(1+S)}\frac{1}{\pi a_0^3}\left(e^{-r_A/a_0} + e^{-r_B/a_0}\right)^2 \\ &= \frac{1}{2(1+S)}\frac{1}{\pi a_0^3} \\ &= (0.561)^2 \times \left(\frac{1}{\pi \times (52.9\text{ pm})^3}\right) = 6.7672 \times 10^{-7}\text{ pm}^{-3}\end{aligned}$$

(Note: The value of C^2 is obtained from equation 9.3.12 and the value of S is from Equation 9.3.12)

To find the density at halfway between A and B for the bonding orbital evalulate at 1/2 R,

$$r_A = \frac{1}{2}R \text{ and } r_B = \frac{1}{2}R$$

$$\psi_+^2 = 3.7 \times 10^{-7}\text{ pm}^{-3}$$

Now multiply this answer by e to generate the electronic charge density,

$$P = \left(3.7 \times 10^{-7}\text{ pm}^{-3}\right) \times \left(1.602 \times 10^{-19}\right)\text{ C} = 5.93 \times 10^{-26} C/pm^3$$

b: Antibonding Molecular Orbital

To find the halfway between A and B for the anti-bonding orbital at $r_A = \frac{1}{2}R$ and $r_B = \frac{1}{2}R$. However, simple inspection of Equation 9.3.10 or Figure 9.3.1 shows this will be zero.

$$e\psi_-^2 = 0$$

Obviously, when multiplied by e, the answer is still zero. This makes logical sense as there would be no charge density between the two atoms as there lies a node with no probability of finding an electron.

Exercise 9.3.1

> Show that for two arbitrary functions $\langle\varphi_B|\varphi_A\rangle$ is the complex conjugate of $\langle\varphi_A|\varphi_B\rangle$ and that these two integrals are equal if the functions are real.

Contributors and Attributions

- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")

9.3: The Overlap Integral is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **9.3: The Overlap Integral** is licensed CC BY-NC-SA 4.0.

9.3B: Evaluating the Overlap Integral (Optional)

The overlap integral is critical to finding the energy in the LCAO-MO approach. For the ground state of the H_2^+ molecule we stated the end result was:

$$S(R) = \langle 1s_A | 1s_B \rangle = e^{-R/a_o} \left(1 + \frac{R}{a_o} + \frac{R^2}{3a_o^2}\right) \tag{9.3B.1}$$

The difficulty in evaluating this integral is a result of the separate spherical coordinate system we use to describe each atom. Thus when we combine the two atoms to make a molecule we end up with three different distances:

- the internuclear separation R and
- the electron position relative to nucleus r_a
- the electron position relative to nucleus r_b

and four angles ϕ_A, θ_A, ϕ_B, and θ_B two from each nucleus.

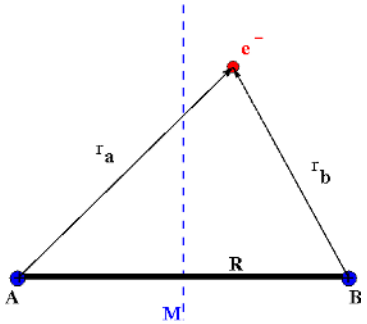


Figure 9.3B.1: Hydrogen molecular ion H_2^+ with fixed nuclei A and B, internuclear distance R.

There are two ways to address the coordinate system of the overlap integral.

1. Change from two spherical coordinates into one elliptic coordinate system.
2. Use the law of cosines to express r_b in terms of the coordinates of atom A.

The natural coordinates for a two-center system like H_2 are elliptical and provides the easiest path to evaluating the integral, however many students are not aware of this method because elliptical coordinates are rarely, maybe never, taught in other math and physics classes. Using the law of cosines is a long and arduous integral as it requires trig and u substitution to complete. A step by step guide to method 1 is provided and an outline for how to use method 2.

Method 1: Two-Center Elliptical Coordinates

In the two-center elliptical system the any point P can be expressed by three coordinates λ, μ, ϕ where:

- $1 \leq \lambda < \infty$
- $-1 \leq \mu \leq 1$
- $0 \leq \theta \leq 2\pi$

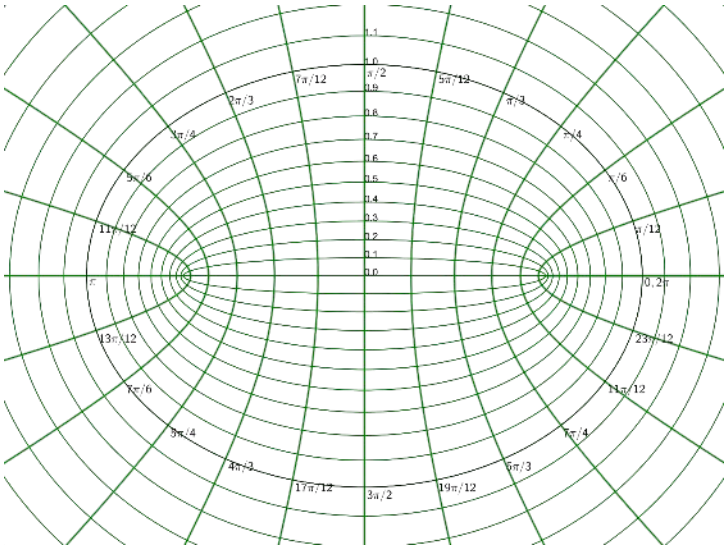


Figure **9.3*B*.1**: Grid for elliptical coordinates. Elliptical grid lines are in even steps of μ of 0.1 units. Hyperbolic grid lines are in even steps of ν of pi/12 for the interval [0,2pi]. a is equal to 1. Image used with permission (GDLF; SharkD).

With this system

$$\lambda = \frac{r_a + r_b}{R} \tag{9.3B.2}$$

$$\mu = \frac{r_a - r_b}{R} \tag{9.3B.3}$$

and the Jacobian or volume element is:

$$d\tau = \frac{R^3}{8}(\lambda^2 - \mu^2)d\lambda d\mu d\theta \tag{9.3B.4}$$

Back to the overlap integral S

$$S = \int 1s_A 1s_B d\tau \tag{9.3B.5}$$

The 1s wavefunction is

$$\psi_{100} = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-Zr/a_0} \tag{9.3B.6}$$

Putting the $1s_A$ and $1s_B$ wave functions into the overlap integral still in two sets spherical coordinates we have:

$$S = \int 1s_A 1s_B d\tau = \int \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-Zr_a/a_0} \times \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-Zr_b/a_0} d\tau \tag{9.3B.7}$$

Still in two sets of spherical coordinates:

$$S = \int 1s_A 1s_B d\tau = \frac{Z^3}{a_0^3\pi}\int e^{-Zr_a/a_0} e^{-Zr_b/a_0} d\tau \tag{9.3B.8}$$

$$S = \int 1s_A 1s_B d\tau = \frac{Z^3}{a_0^3\pi}\int e^{\frac{-Z(r_a+r_b)}{a_0}} d\tau \tag{9.3B.9}$$

Notice that we only have the r_a and r_b coordinates represented. The R distance in the integral is not present.

Change coordinates:

From our definition above

$$R\lambda = r_a + r_b \tag{9.3B.10}$$

$$S = \int 1s_A 1s_B d\tau = \frac{Z^3}{a_0^3 \pi} \int e^{\frac{-RZ\lambda}{a_0}} d\tau \tag{9.3B.11}$$

Look how R becomes part of the integral naturally when the coordinate system changes. There is no need to convert the spherical angles ϕ_A, θ_A, ϕ_B, and θ_B into elliptical coordinates because they are not present in the original spherical integral, or in the wavefunctions.

Put in the Jacobian and:

$$S = \int 1s_A 1s_B d\tau = \frac{Z^3}{a_0^3 \pi} \frac{R^3}{8} \int_1^\infty \int_{-1}^1 \int_0^{2\pi} e^{\frac{-RZ\lambda}{a_0}} (\lambda^2 - \mu^2) d\lambda d\mu d\theta \tag{9.3B.12}$$

$$S = \int 1s_A 1s_B d\tau \tag{9.3B.13}$$

$$= \frac{Z^3}{a_0^3 \pi} \frac{R^3}{8} \int_0^{2\pi} d\theta \left[\int_1^\infty \int_{-1}^1 \lambda^2 e^{\frac{-RZ\lambda}{a_0}} d\lambda d\mu - \int_1^\infty \int_{-1}^1 \mu^2 e^{\frac{-RZ\lambda}{a_0}} d\lambda d\mu \right]$$

Lets simplfy this a little bit by noting that $Z = 1$ for hydrogen and evaluating the angular part.

$$S = \int 1s_A 1s_B d\tau = \frac{1}{a_0^3} \frac{R^3}{4} \left[\int_1^\infty \int_{-1}^1 \lambda^2 e^{\frac{-R\lambda}{a_0}} d\lambda d\mu - \int_1^\infty \int_{-1}^1 \mu^2 e^{\frac{-R\lambda}{a_0}} d\lambda d\mu \right] \tag{9.3B.14}$$

Next the $d\mu$ integrals

$$S = \int 1s_A 1s_B d\tau = \frac{1}{a_0^3} \frac{R^3}{4} \left[2 \int_1^\infty \lambda^2 e^{\frac{-R\lambda}{a_0}} d\lambda - \frac{2}{3} \int_1^\infty e^{\frac{-R\lambda}{a_0}} d\lambda \right] \tag{9.3B.15}$$

To finish we need the excellent integral

$$\int_t^\infty z^n e^{-bz} = \frac{n!}{b^{n+1}} \left(1 + bt + \frac{b^2 t^2}{2!} + ... \frac{b^n t^n}{n!} \right) \tag{9.3B.16}$$

where $n = 0, 1, 2, 3, 4, ...$ and $b > 0$

In our case $n = 0$ and 2 with $t = 1$ and $b = R/a_0$

$$S = \int 1s_A 1s_B d\tau = \frac{1}{a_0^3} \frac{R^3}{4} \left[2 \frac{2a_0^3}{R^3} e^{-R/a_0} \left(1 + \frac{R}{a_0} + \frac{R^2}{2a_0^2} \right) - \frac{2}{3} \frac{a_0}{R} e^{-R/a_0} (1) \right] \tag{9.3B.17}$$

Distribute the constant out front

$$S = \int 1s_A 1s_B d\tau = \left[e^{-R/a_0} \left(1 + \frac{R}{a_0} + \frac{R^2}{2a_0^2} \right) - \frac{R^2}{6a_0^2} e^{-R/a_0} \right] \tag{9.3B.18}$$

Pull out the exponent and ....

$$S = \int 1s_A 1s_B d\tau = e^{-R/a_0} \left[1 + \frac{R}{a_0} + \frac{R^2}{2a_0^2} - \frac{R^2}{6a_0^2} \right] \tag{9.3B.19}$$

The Answer -

$$S(R) = \int 1s_A 1s_B d\tau = e^{-R/a_0} \left[1 + \frac{R}{a_0} + \frac{R^2}{3a_0^2} \right] \tag{9.3B.20}$$

Method 2: The Law of Cosines

Work in Progress...

References

- http://www.colby.edu/chemistry/PChem...structions.pd
- Recursion formulae for calculation of overlap integrals, January 1987. 10.1002/jcc.540080102

- Formulas and Numerical Tables for Overlap Integrals R. S. Mulliken, C. A. Rieke, D. Orloff and H. Orloff, J. Chem. Phys. 17, 1248 (1949); http://dx.doi.org/10.1063/1.1747150

9.3B: Evaluating the Overlap Integral (Optional) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.4: Chemical Bond Stability

Learning Objectives

- Identify the nature of the energy of molecular orbitals of a diatomic as a function of intermolecular distance
- Identify the three integrals involved in calculation the total Molecular Orbital Energy: coulomb Integral, exchange integral, and overlap integral

As shown previously, we can construct two molecular orbitals for the H_2^+ system using the LCAO approximation with a basis set of two 1s atomic orbitals (i.e., the $1s$ orbitals on hydrogen A ($1s_A$) and hydrogen B ($1s_B$):

$$|\psi_{\pm}\rangle = \frac{1}{\sqrt{2(1 \pm S)}}(1s_A \pm 1s_B) \tag{9.4.1}$$

The energy of these two molecular orbitals can be calculated from the expectation value integral of the Hamiltonian,

$$E_{\pm} = \left\langle \psi_{\pm} | \hat{H}_{elec} | \psi_{\pm} \right\rangle \tag{9.4.2}$$

which can be expanded using the expanded molecular orbital wavefunctions in Equations 9.4.1 to give

$$E_{\pm} = \frac{1}{2(1 \pm S)} \tag{9.4.3}$$

$$\left[\underbrace{\left\langle 1s_A | \hat{H}_{elec} | 1s_A \right\rangle}_{H_{AA}} + \underbrace{\left\langle 1s_B | \hat{H}_{elec} | 1s_B \right\rangle}_{H_{BB}} \pm \underbrace{\left\langle 1s_A | \hat{H}_{elec} | 1s_B \right\rangle}_{H_{AB}} \pm \underbrace{\left\langle 1s_B | \hat{H}_{elec} | 1s_A \right\rangle}_{H_{BA}} \right]$$

where S is the overlap integral between the two atomic orbitals of the basis. The four integrals in Equation 9.4.3 can be represented by H_{AA}, H_{BB}, H_{AB}, and H_{BA}, respectively.

$$E_{\pm} = \frac{1}{2(1 \pm S)} \left[H_{AA} + H_{BB} \pm H_{AB} \pm H_{BA} \right] \tag{9.4.4}$$

Exercise 9.4.1

Show that Equation 9.4.2 expands to give Equation 9.4.3 within the LCAO approximation that uses a basis set of only two 1s atomic orbitals.

Notice that A and B appear equivalently in the Hamiltonian operator for H_2^+. This equivalence means that integrals involving $1s_A$ must be the same as corresponding integrals involving $1s_B$, i.e.

$$H_{AA} = H_{BB} \tag{9.4.5}$$

and since the wavefunctions are real

$$|A\rangle = \langle A|$$

so

$$H_{AB} = H_{BA} \tag{9.4.6}$$

These two equalities simplify Equation 9.4.4:

$$E_{\pm} = \frac{1}{1 \pm S}(H_{AA} \pm H_{AB}) \tag{9.4.7}$$

Now examine the details of H_{AA} after inserting the Hamiltonian operator for H_2^+ (Equation 9.2.1):

$$H_{AA} = \underbrace{\left\langle 1s_A \left| -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0 r_A} \right| 1s_A \right\rangle}_{E_H} + \frac{e^2}{4\pi\epsilon_0 R}\cancelto{1}{\langle 1s_A | 1s_A \rangle} - \underbrace{\left\langle 1s_A \left| \frac{e^2}{4\pi\epsilon_0 r_B} \right| 1s_A \right\rangle}_{J_{AB}} \tag{9.4.8}$$

- The first term is just the integral for the energy of the hydrogen atom of the 1s orbital, E_H.
- The second integral is equal to 1 by normalization; the prefactor is just the Coulomb repulsion of the two protons.
- The last integral, *including the minus sign*, is represented by J and is called the **Coulomb integral**.

Physically J_{AB} is the potential energy of interaction of the electron located around proton A with proton B. It is negative because it is an attractive interaction. It is the average interaction energy of an electron described by the $1s_A$ function with proton B.

J The Coulomb Integral ()

The **Coulomb Integral** is the potential energy of electrostatic repulsion between the electron with the electron density in $1S_A$ and the the electron with the electron density function $1S_B$

Now consider H_{AB}.

$$\underbrace{= \left\langle 1s_A \left| -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0 r_B} \right| 1s_B \right\rangle}_{E_H S} + \frac{e^2}{4\pi\epsilon_0 R}\cancelto{S}{\langle 1s_A | 1s_B \rangle} - \underbrace{\left\langle 1s_A \left| \frac{e^2}{4\pi\epsilon_0 r_A} \right| 1s_B \right\rangle}_{K_{AB}} \tag{9.4.9}}$$

- In the first integral we have the hydrogen atom Hamiltonian and the H atom function $1s_B$. The function $1s_B$ is an eigenfunction of the operator with eigenvalue E_H. Since E_H is a constant it factors out of the integral, which then becomes the overlap integral, S. The first integral therefore reduces to $E_H S$.
- The second term is just the Coulombic energy of the two protons times the overlap integral.
- The third term, *including the minus sign*, is given the symbol K and is called the *exchange integral* because the electron is described by the $1s_A$ orbital on one side and by the $1s_B$ orbital on the other side of the operator. The electron changes or exchanges position in the molecule.

K The Exchange Integral ()

In a **coulomb integral**, the electron always is in the same orbital; whereas, in an *Exchange Integral*, the electron is in one orbital on one side of the operator and in a different orbital on the other side.

Using the expressions for H_{AA} (Equation 9.4.8) and H_{AB} (Equation 9.4.9) and substituting into Equation 9.4.7 produces:

$$E_{\pm} = \frac{1}{1 \pm S}\left[\left(E_H + \frac{e^2}{4\pi\epsilon_0 R}\right)(1 \pm S) + J \pm K\right] \tag{9.4.10}$$

$$= E_H + \frac{e^2}{4\pi\epsilon_0 R} + \frac{J \pm K}{1 \pm S} \tag{9.4.11}$$

Equation 9.4.11 tells us that the energy of the H_2^+ molecule is the energy of a hydrogen atom plus the repulsive energy of two protons plus some additional electrostatic interactions of the electron with the protons. These additional interactions are given by the last term

$$\frac{J \pm K}{1 \pm S}.$$

If the protons are infinitely far apart then only E_H is nonzero, which we can set to zero by subtracting off:

$$\Delta E_{\pm} = E_{\pm} - E_H \tag{9.4.12}$$

$$= \frac{e^2}{4\pi\epsilon_0 R} + \frac{J \pm K}{1 \pm S} \tag{9.4.13}$$

To get a chemical bond and a stable H_2^+ molecule $\Delta E_{\pm}$ must be less than zero and have a minimum, i.e. $\frac{J \pm K}{1 \pm S}$ must be sufficiently negative to overcome the positive repulsive energy of the two protons $\frac{e^2}{4\pi\epsilon_0 R}$ for some value of R. For large R, these terms are zero, and for small R, the **Coulomb repulsion of the protons rises to infinity.**

? Exercise 9.4.2

Show that Equation 9.2.1 follows from Equation 9.4.7

The Coulomb and Exchange Integrals

We will examine more closely how the Coulomb repulsion term and the integrals J, K, and S depend on the separation of the protons, but first we want to discuss the physical significance of J, the Coulomb integral, and K, the exchange integral. J and K have been defined as

$$J_{AB} = \left\langle 1s_A \left| \frac{-e^2}{4\pi\epsilon_0 r_B} \right| 1s_A \right\rangle = -\int \varphi^*_{1s_A}(r)\varphi_{1s_A}(r)\frac{e^2}{4\pi\epsilon_0 r_B}d\tau \tag{9.4.14}$$

$$K_{AB} = \left\langle 1s_A \left| \frac{-e^2}{4\pi\epsilon_0 r_A} \right| 1s_B \right\rangle = -\int \varphi^*_{1s_A}(r)\varphi_{1s_B}(r)\frac{e^2}{4\pi\epsilon_0 r_A}d\tau \tag{9.4.15}$$

Figure 9.4.2 shows graphs of the four terms contributing to the energy of H_2^+ (Equation 9.4.13). In Figure 9.4.2 , you can see that as the internuclear distance R approaches zero,

- the Coulomb repulsion of the two protons goes from near zero to a large positive number,
- the overlap integral goes for zero to one, and
- J and K become increasingly negative.

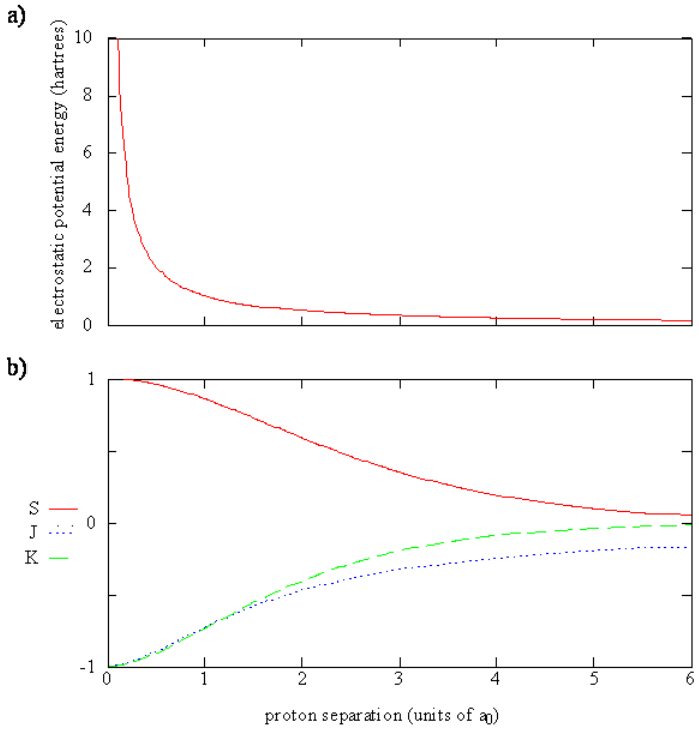


Figure 9.4.2 : (a) The electrostatic energy (in hartrees, 27.2 eV) of two protons separated by a distance R in units of the Bohr radius (52.92 pm). (b) The overlap (S), Coulomb (J), and exchange (K) integrals at different proton separations. The units for J and K are hartrees; S has no units. (CC BY-SA-NC; David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski via ChemEdDL)

Note that both J and K integrals are **negative** since all quantities in the integrands of Equation 9.4.14 and 9.4.15 are positive. In the Coulomb integral, $e\varphi^*_{1s_A}(r)\varphi_{1s_A}(r)$ is the charge density of the electron around proton A, since **r** represents the coordinates of the

electron relative to proton A. Since r_B is the distance of this electron to proton B, the Coulomb integral gives the potential energy of the charge density around proton A interacting with proton B. J can be interpreted as an average potential energy of this interaction because $e\varphi^*_{1s_A}(r)\varphi_{1a_A}(r)$ is the probability density for the electron at point r, and $\frac{e^2}{4\pi\epsilon_0 r_B}$ is the potential energy of the electron at that point due to the interaction with proton B. Essentially, J accounts for the attraction of proton B to the electron density of hydrogen atom A. As the two protons get further apart, this integral goes to zero because all values for r_B become very large and all values for $1/r_B$ become very small.

In the **exchange integral**, K, the product of the two functions is nonzero **only** in the regions of space where the two functions overlap. If one function is zero or very small at some point then the product will be zero or small. The exchange integral also approaches zero as internuclear distances increase because the both the overlap and the $1/r$ values become zero. The product $e\varphi^*_{1s_A}(r)\varphi_{1a_B}(r)$ is called the *overlap charge density*. Since the overlap charge density is significant in the region of space between the two nuclei, it makes an important contribution to the chemical bond. The exchange integral, K, is the potential energy due to the interaction of the overlap charge density with one of the protons. While J accounts for the attraction of proton B to the electron density of hydrogen atom A, K accounts for the added attraction of the proton due the build-up of electron charge density between the two protons.

? Exercise 9.4.3

Write a paragraph describing in your own words the physical significance of the Coulomb and exchange integrals for H_2^+.

Figure 9.4.3 shows the energy of H_2^+ relative to the energy of a separated hydrogen atom and a proton as given by Equation 9.4.11. For the electron in the ψ_- orbital, the energy of the molecule, $E_{el}(R)$, **always** is greater than the energy of the separated atom and proton.

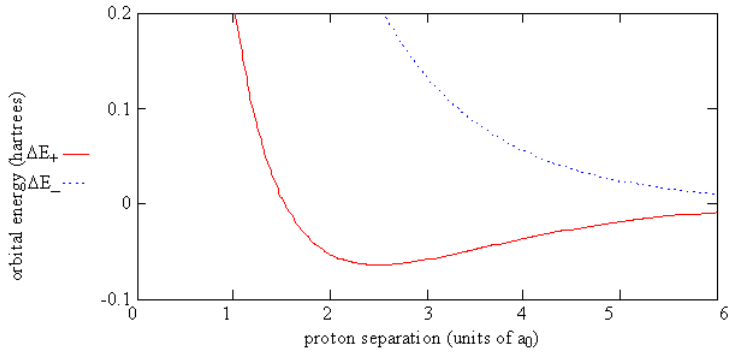


Figure 9.4.3 : Energy of the H_2^+ bonding molecular orbital ΔE_+ and the molecular orbital ΔE_-, relative to the energy of a separated hydrogen atom and proton.

For the electron in the ψ_+ orbital, you can see that the big effect for the energy of the bonding orbital, $E_+(R)$, is the balance between the repulsion of the two protons $\frac{e^2}{4\pi\epsilon_0 R}$ and J and K, which are both negative. J and K manage to compensate for the repulsion of the two protons until their separation is less than 100 pm (i.e the energy is negative up until this point), and a minimum in the energy is produced at 134 pm. This minimum represents the formation of a chemical bond. The effect of S is small. It only causes the denominator in Equation 9.4.11 to increase from 1 to 2 as R approaches 0.

For the antibonding orbital, $-K$ is a positive quantity and essentially cancels J so there is not sufficient compensation for the Coulomb repulsion of the protons. The effect of the -K in the expression, Equation 9.4.11, for E_- is to account for the absence of overlap charge density and the enhanced repulsion because the charge density between the protons for ψ_- is even lower than that given by the atomic orbitals.

Contributors and Attributions

- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")

9.4: Chemical Bond Stability is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **9.4: Chemical Bond Stability** is licensed CC BY-NC-SA 4.0.

9.5: Bonding and Antibonding Orbitals

Learning Objectives

- Characterize the bonding and anti-bonding molecular orbitals in $\mathbf{H^+}$

The two molecular orbitals of the $\mathbf{H^+}$ ion were created via the linear combinations of atomic orbitals (LCAOs) approximation were created from the sum and the difference of two atomic orbitals. Within this approximation, the jth molecular orbital can be expressed as a linear combination of many atomic orbitals $\{\phi_i\}$:

$$|\psi_J\rangle = \sum_i^N c_{J,i}|\phi_i\rangle \tag{9.5.1}$$

A molecule will have as many molecular orbitals as there are atomic orbitals used in the basis set (N in Equation 9.5.1). Adding two atomic orbitals corresponds to *constructive* interference between two waves, thus reinforcing their intensity; the internuclear electron probability density is *increased*. The molecular orbital corresponding to the sum of the two H 1*s* orbitals is called a σ_{1s} combination (parts (a) and (b) of Figure 9.5.1).

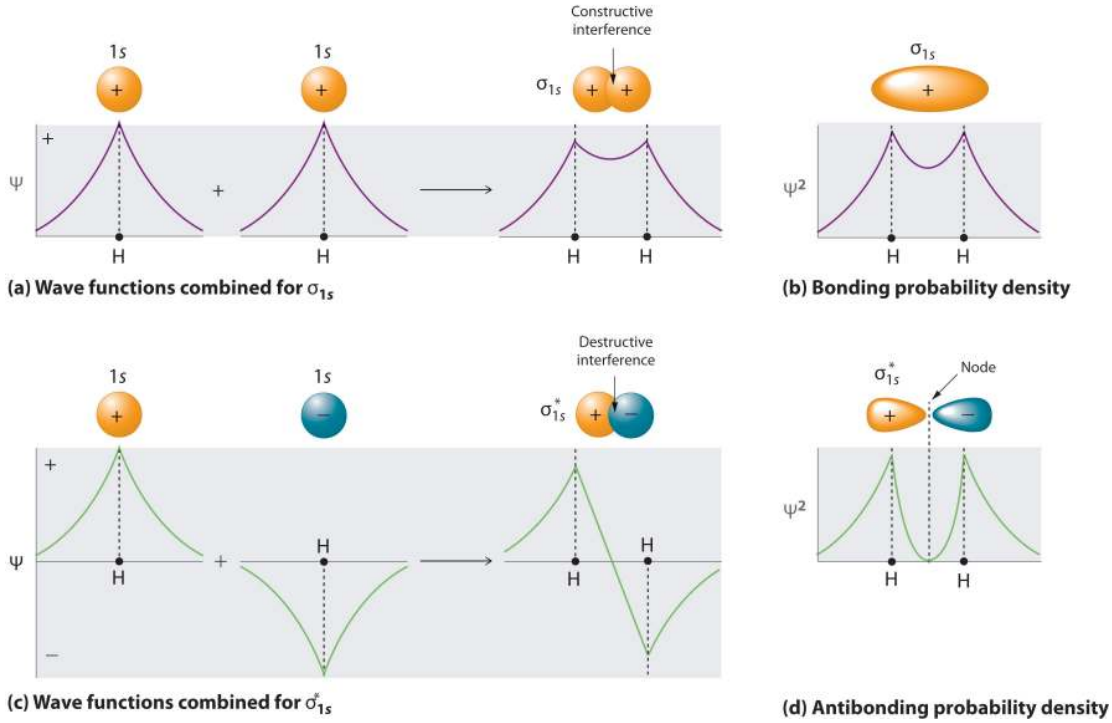


Figure 9.5.1 : Molecular Orbitals for the H_2 Molecule. (a) This diagram shows the formation of a bonding σ_{1s} molecular orbital for H_2 as the sum of the wavefunctions (Ψ) of two H 1s atomic orbitals. (b) This plot of the square of the wavefunction (Ψ^2) for the bonding σ1s molecular orbital illustrates the increased electron probability density between the two hydrogen nuclei. (Recall that the probability density is proportional to the square of the wavefunction.) (c) This diagram shows the formation of an antibonding σ^*_{1s} molecular orbital for H_2 as the difference of the atomic orbital wavefunctions (Ψ) of two H 1s atomic orbitals. (d) This plot of the square of the wavefunction (Ψ2) for the σ^*_{1s} antibonding molecular orbital illustrates the node corresponding to zero electron probability density between the two hydrogen nuclei. (CC BY-SA-NC; anonymous).

In the sigma (σ) orbital, the electron density along the internuclear axis and between the nuclei has *cylindrical symmetry*; that is, all cross-sections perpendicular to the internuclear axis are circles. The subscript 1*s* denotes the atomic orbitals from which the molecular orbital was derived.

$$|\sigma_{1s}\rangle = \frac{1}{\sqrt{2(1+S)}}\left(|1s_A\rangle + |1s_B\rangle\right) \tag{9.5.2}$$

Conversely, subtracting one atomic orbital from another corresponds to *destructive* interference between two waves, which reduces their intensity and causes a *decrease* in the internuclear electron probability density (part (c) and part (d) in Figure 9.5.1). The resulting pattern contains a *node* where the electron density is zero. The molecular orbital corresponding to the difference is called σ^*_{1s} and has a region of zero electron probability, a nodal plane, perpendicular to the internuclear axis:

$$|\sigma^*_{1s}\rangle = \frac{1}{\sqrt{2(1-S)}}\left(|1s_A\rangle - |1s_B\rangle\right) \tag{9.5.3}$$

The electron density in the σ_{1s} molecular orbital is greatest between the two positively charged nuclei, and the resulting electron–nucleus electrostatic attractions reduce repulsions between the nuclei. Thus the σ_{1s} orbital represents a bonding molecular orbital. A molecular orbital that forms when atomic orbitals or orbital lobes with the same sign interact to give increased electron probability between the nuclei due to constructive reinforcement of the wavefunctions. In contrast, electrons in the σ^*_{1s} orbital are generally found in the space outside the internuclear region. Because this allows the positively charged nuclei to repel one another, the σ^*_{1s} orbital is an antibonding molecular orbital (a molecular orbital that forms when atomic orbitals or orbital lobes of opposite sign interact to give decreased electron probability between the nuclei due to destructive reinforcement of the wavefunctions).

> *Antibonding orbitals contain a node perpendicular to the internuclear axis; bonding orbitals do not.*

Because electrons in the σ_{1s} orbital interact simultaneously with both nuclei, they have a lower energy than electrons that interact with only one nucleus. This means that the σ_{1s} molecular orbital has a *lower* energy than either of the hydrogen 1*s* atomic orbitals. Conversely, electrons in the σ^*_{1s} orbital interact with only one hydrogen nucleus at a time. In addition, they are farther away from the nucleus than they were in the parent hydrogen 1*s* atomic orbitals. Consequently, the σ^*_{1s} molecular orbital has a *higher* energy than either of the hydrogen 1*s* atomic orbitals. The σ_{1s} (bonding) molecular orbital is *stabilized* relative to the 1*s* atomic orbitals, and the σ^*_{1s} (antibonding) molecular orbital is *destabilized*. The relative energy levels of these orbitals are shown in the energy-level diagram (a schematic drawing that compares the energies of the molecular orbitals (bonding, antibonding, and nonbonding) with the energies of the parent atomic orbitals) in Figure 9.5.2

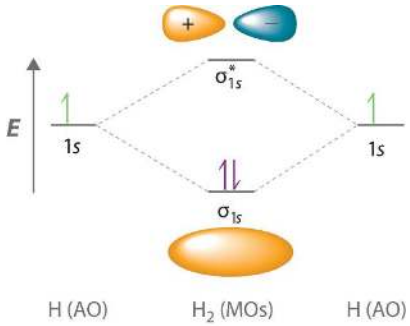


Figure 9.5.2 : Molecular Orbital Energy-Level Diagram for H2. The two available electrons (one from each H atom) in this diagram fill the bonding σ1s molecular orbital. Because the energy of the σ1s molecular orbital is lower than that of the two H 1s atomic orbitals, the H2 molecule is more stable (at a lower energy) than the two isolated H atoms.

> *A bonding molecular orbital is always lower in energy (more stable) than the component atomic orbitals, whereas an antibonding molecular orbital is always higher in energy (less stable).*

Expanding Beyond the 1s Orbital Basis Set

This picture of bonding in H_2^+ in the previous section is very simple, but gives reasonable results when compared to an exact calculation. The equilibrium bond distance is 134 pm compared to 106 pm (exact), and a dissociation energy is 1.8 eV compared to 2.8 eV (exact). To better describe chemical bonding we need to account for the increase in electron density between the two nuclei. The 1s orbitals alone are not particularly good for this purpose because they are spherically symmetric and show no preference for the space between the atomic nuclei. The use of additional atomic orbitals can correct this situation and provide additional parameters, which can be optimized by the linear variational method, to give a better function with a lower energy and more accurate description of the charge density.

The energy of the non-normalized molecular orbital can be calculated from the expectation value integral of the Hamiltonian,

$$E_J = \frac{\left\langle \psi_J | \hat{H}_{elec} | \psi_J \right\rangle}{\langle \psi_J | \psi_J \rangle} \tag{9.5.4}$$

This is the variational energy using $|\psi_J\rangle$ as the trail wavefunction. After substituting the LCAO expansion for $|\psi_J\rangle$ (Equation 9.5.1) into the energy expression of Equation 9.5.4 results in:

$$E_J = \frac{\left\langle \sum_i c_{J,i}^* \phi_i \middle| \hat{H}_{elec} \middle| \sum_j c_{J,i} \phi_j \right\rangle}{\left\langle \sum_i c_{J,i}^* \phi_j \middle| \sum_j c_{J,j} \phi_j \right\rangle} \tag{9.5.5}$$

$$= \frac{\sum_{i,j} c_{J,i}^* c_{J,j} \langle \phi_i | \hat{H}_{elec} | \phi_j \rangle}{\sum_{i,j} c_{J,i}^* c_{J,j} \langle \phi_i | \phi_j \rangle} \tag{9.5.6}$$

$$= \frac{\sum_{i,j} c_{J,i}^* c_{J,j} H_{ij}}{\sum_{i,j} c_{J,i}^* c_{J,j} S_{ij}} \tag{9.5.7}$$

where H_{ij} is the Hamiltonian matrix element.

$$H_{ij} = \langle \phi_i | \hat{H}_{elec} | \phi_j \rangle$$

Following the variational theorem, to determine the coefficients of the LCAO expansion c_i, we need to minimize E_J

$$\frac{\partial E_J}{\partial c_k} = 0 \tag{9.5.8}$$

for all k. This requires solving N linear equations to hold true (where N is the number of atomic orbitals in the basis)

$$\sum_{i=1}^{N} c_i (H_{ki} - E S_{ki}) = 0 \tag{9.5.9}$$

These equations are the **secular equations** and were discussed previously in the context of the linear variational method approximation. For the two basis set expansion (N) in Figure 9.5.1 , these are

$$\begin{aligned} c_1(H_{11} - ES_{11}) + c_2(H_{12} - ES_{12}) &= 0 \\ c_1(H_{12} - ES_{12}) + c_2(H_{22} - ES_{22}) &= 0 \end{aligned} \tag{9.5.10}$$

where c_1 and c_2 are the coefficients in the linear combination of the atomic orbitals used to construct the molecular orbital. Writing this set of homogeneous linear equations in matrix form gives

$$\begin{pmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} \\ H_{12} - ES_{12} & H_{22} - ES_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{9.5.11}$$

Solving these secular equations with N different atomic orbitals in the expansion (Equation 9.5.1) requires finding the N roots of an N order polynomial.

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} & \ldots \\ H_{12} - ES_{12} & H_{22} - ES_{22} & \ldots \\ \ldots & \ldots & \ldots \end{vmatrix} = 0 \tag{9.5.12}$$

Each molecular orbital ($|\psi_J\rangle$) from this treatment has an energy E_J that is given by a different set of coefficients, $\{c_{ij}\}$ where i runs over all N functions in the basis (i.e., number of the atomic orbitals in the LCAO approximation of Equation 9.5.1), and J runs over molecular orbitals. Solve the set of linear equations using that specific E_J to determine c_{ij} values.

Steps in a Solving the Secular Equations

1. Select a set of N basis functions
2. Determine all N(N –1)/2 values of both H_{ij} and S_{ij}
3. Form the secular determinant; determine N roots E_j of secular equation
4. For each E_J solve the set of linear equations to determine the basis set coefficients \(c_{ij\}) for the j-th molecular orbital

For more information on solving the Secular equations check here.

The greater the number of atomic orbitals N that combine to genera the molecular orbitals (Equation 9.5.1), the more accurate the LCAO approximation is. This is expected based on our discussions of the variational method examples. Hence, the ψ_+ and ψ_- molecular orbitals for H_2^+ are better expressed with higher energy hydrogenic wavefunctions

$$|\psi_J\rangle = c_{J,1}1s_A + c_{J,2}1s_B + c_{J,3}2s_A + c_{J,4}2s_B + c_{J,5}2p_{z,A} + c_{J,6}2p_{z,B} \tag{9.5.13}$$

The reasons that only the p_z atomic orbitals are included in this expansion are discussed later.

Contributors and Attributions

- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")

9.5: Bonding and Antibonding Orbitals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **9.5: Bonding and Antibonding Orbitals** is licensed CC BY-NC-SA 4.0.

9.6: A Simple Molecular-Orbital Treatment of H_2 Places Both Electrons in a Bonding Orbital

To describe the electronic states of molecules, we construct wavefunctions for the electronic states by using molecular orbitals. These wavefunctions are approximate solutions to the Schrödinger equation. A mathematical function for a molecular orbital is constructed, ψ_i, as a linear combination of other functions, φ_j, which are called basis functions because they provide the basis for representing the molecular orbital.

$$|\psi_i\rangle = \sum_j c_{ij}\varphi_j \tag{9.6.1}$$

where

- j is the index for the jth basis function (e.g., atomic orbital)
- i is the ith molecular orbitals and
- c_{ij} is the expansion coefficient of the jth basis function for the ith molecular orbital.

The variational method is used to find values for parameters in the basis functions and for the constant coefficients in the linear combination that optimize these functions, i.e. make them as good as possible. The criterion for quality in the variational method is making the ground state energy of the molecule as low as possible. Here and in the rest of this chapter, the following notation is used: σ is a general spin function (can be either α or β), φ is the basis function (this usually represents an atomic orbital), ψ is a molecular orbital, and Ψ is the electronic state wavefunction (representing a single Slater determinant or linear combination of Slater determinants).

The ultimate goal is a mathematical description of electrons in molecules that enables chemists and other scientists to develop a deep understanding of chemical bonding and reactivity, to calculate properties of molecules, and to make predictions based on these calculations. Just as for atoms, each electron in a molecule can be described by a product of spin-orbitals. Since electrons are fermions, the electronic wavefunction must be **antisymmetric** with respect to the permutation of any two electrons. A Slater determinant containing the molecular spin orbitals produces the antisymmetric wavefunction. For example for two electrons,

$$\Psi(r_1, r_2) = \frac{1}{\sqrt{2}}\begin{vmatrix}\psi_A(r_1)\alpha(1) & \psi_B(r_1)\beta(1)\\ \psi_A(r_2)\alpha(2) & \psi_B(r_2)\beta(2)\end{vmatrix} \tag{9.6.2}$$

Solving the Schrödinger equation in the orbital approximation will produce a set of spatial molecular orbitals, each with a specific energy, ϵ. Following the Aufbau Principle, two electrons with different spins (α and β, consistent with the Pauli Exclusion Principle discussed for muliti-electron atoms) are assigned to each spatial molecular orbital in order of increasing energy. For the ground state of the 2n electron molecule, the n lowest energy spatial orbitals will be occupied, and the electron configuration will be given as $\psi_1^2\psi_2^2\psi_3^2 \ldots \psi_n^2$. The electron configuration also can be specified by an orbital energy level diagram as shown in Figure 9.6.1 . Higher energy configurations exist as well, and these configurations produce excited states of molecules. Some examples are shown in Figure 9.6.1 .

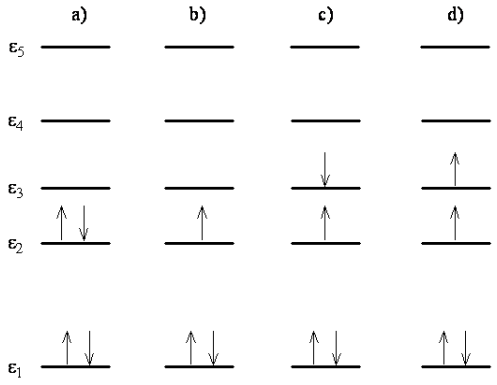


Figure 9.6.1 : a) The lowest energy configuration of a closed-shell system. b) The lowest energy configuration of an open-shell radical. c) An excited singlet configuration. d) An excited triplet configuration.

Symmetry

Molecular orbitals usually are identified by their symmetry or angular momentum properties. For example, a typical symbol used to represent an orbital in an electronic configuration of a diatomic molecule is $2\sigma_g^2$. The superscript in symbol means that this orbital is occupied by two electrons; the prefix means that it is the second sigma orbital with *gerade* symmetry.

Diatomic molecules retain a component of angular momentum along the internuclear axis. The molecular orbitals of diatomic molecule therefore can be identified in terms of this angular momentum. A Greek letter, e.g. σ or π, encodes this information, as well as information about the symmetry of the orbital. A σ means the component of angular momentum is 0, and there is no node in any plane containing the internuclear axis, so the orbital must be symmetric with respect to reflection in such a plane. A π means there is a node and the wavefunction is antisymmetric with respect to reflection in a plane containing the internuclear axis. For homonuclear diatomic molecules, a *g* or a *u* is added as a subscript to designate whether the orbital is symmetric or antisymmetric with respect to the center of inversion of the molecule.

A homonuclear diatomic molecule has a center of inversion in the middle of the bond. This center of inversion means that $\psi(x,y,z) = \pm\psi(-x,-y,-z)$ with the origin at the inversion center. Inversion takes you from (x,y,z) to $(-x,-y,-z)$. For a heteronuclear diatomic molecule, there is no center of inversion so the symbols g and u are not used. A prefix 1, 2, 3, etc. simply means the first, second, third, etc. orbital of that type. We can specify an electronic configuration of a diatomic molecule by these symbols by using a superscript to denote the number of electrons in that orbital, e.g. the lowest energy configuration of N_2 is

$$1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 1\pi_u^4 3\sigma_g^2 \tag{9.6.3}$$

Contrasting LCAO with other Quantum Chemistry Methods

As we have already seen, the LCAO approach is very approximate, yielding only qualitative results. It should be noted that the Hartree-Fock method discussed earlier for atoms can also be used for molecules. For example, the molecule He_2^+ has three electrons, and the Li atom also has three electrons. As usual with Hartree-Fock, the idea is to optimize the shapes of the single-electron orbitals $\psi_1(r)$, $\psi_2(r)$ and $\psi_3(r)$ by minimizing the guess to the ground state energy E_g. Of course, we will not get the same answer as for Li because there is a different V_{en} energy for He_2^+ due to the presence of two positively charge nuclei $(charge = +2e)$ separated by a distance R in contrast to the single $+3e$ charged nucleus for Li. When the shapes of the orbitals are optimized, we also obtain three energy ε_1, ε_2 and ε_3. Not unexpectedly, we find that two of the HF orbitals resemble $1\sigma_g$ while the third resembles $1\sigma_u^*$ and the first two energies ε_1 and ε_2 will be nearly equal, while the third ε_3 will be noticeably higher. To contrast with the LCAO approach, in LCAO, we do not optimize the shapes of the orbitals (these are assumed *a priori* to be $1s$ shaped). All we do is choose the mixing coefficients so as to minimize the guess to the ground-state energy E_g.

We note, finally, that the density functional theory alluded to earlier can also be used for molecules. It is often the case that density functional theory yields a more accurate description than Hartree-Fock, but this depends on the molecule. In any case, both are more accurate than LCAO. There is also a hierarchy of methods called *post Hartree-Fock* methods, all of which are based on the wavefunction rather than the electron density, that can be used to improve upon the HF approximation systematically. The greater the accuracy that is desired for the calculation, the more costly computationally the post HF methods become, so quantum chemistry is often a trade-off between accuracy and efficiency, an issue that becomes more critical to take into account when calculations on large molecules must be carried out!

Contributors and Attributions

- David M. Hanson, Erica Harvey, Robert Sweeney, Theresa Julia Zielinski ("Quantum States of Atoms and Molecules")
- Mark Tuckerman (New York University)

9.6: A Simple Molecular-Orbital Treatment of H₂ Places Both Electrons in a Bonding Orbital is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **9.6: A Simple Molecular-Orbital Treatment of H₂ Places Both Electrons in a Bonding Orbital** is licensed CC BY-NC-SA 4.0.

9.7: Molecular Orbitals Can Be Ordered According to Their Energies

The LCAO-MO method that we used for H_2^+ can be applied qualitatively to homonuclear diatomic molecules to provide additional insight into chemical bonding. A more quantitative approach also is helpful, especially for more complicated situations, like heteronuclear diatomic molecules and polyatomic molecules. When two atoms are close enough for their valence orbitals to overlap significantly, the filled inner electron shells are largely unperturbed; hence they are often ignored in constructing molecular orbitals. This means that we can focus our attention on the molecular orbitals derived from valence atomic orbitals.

Molecular Orbitals Formed from *ns* Orbitals

The molecular orbitals diagrams formatted for the dihydrogen species are similar to the diagrams to any **homonuclear diatomic molecule** with two identical alkali metal atoms (Li_2 and Cs_2, for example) is shown in part (a) in Figure 9.7.1 , where M represents the metal atom. Only two energy levels are important for describing the valence electron molecular orbitals of these species: a σ_{ns} bonding molecular orbital and a σ^*_{ns} antibonding molecular orbital. Because each alkali metal (M) has an ns^1 valence electron configuration, the M_2 molecule has two valence electrons that fill the σ_{ns} bonding orbital. As a result, a bond order of 1 is predicted for all homonuclear diatomic species formed from the alkali metals (Li_2, Na_2, K_2, Rb_2, and Cs_2). The general features of these M_2 diagrams are identical to the diagram for the H_2 molecule. Experimentally, all are found to be stable in the gas phase, and some are even stable in solution.

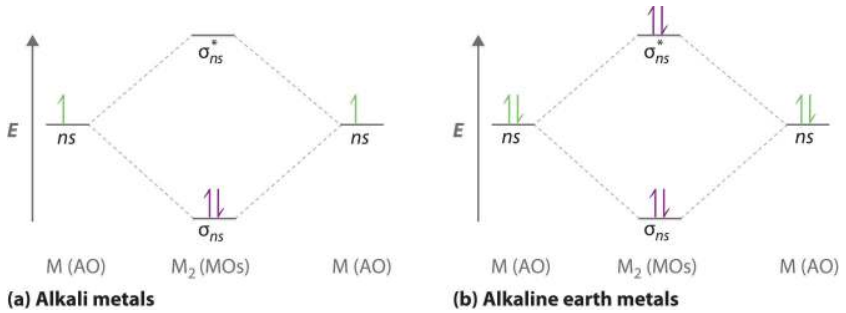


Figure 9.7.1 : Molecular Orbital Energy-Level Diagrams for Alkali Metal and Alkaline Earth Metal Diatomic (M_2) Molecules. (a) For alkali metal diatomic molecules, the two valence electrons are enough to fill the σns (bonding) level, giving a bond order of 1. (b) For alkaline earth metal diatomic molecules, the four valence electrons fill both the σ_{ns} (bonding) and the σ^*_{ns} (nonbonding) levels, leading to a predicted bond order of 0. (CC BY-SA-NC; anonymous).

Similarly, the molecular orbital diagrams for homonuclear diatomic compounds of the alkaline earth metals (such as Be_2), in which each metal atom has an ns^2 valence electron configuration, resemble the diagram for the He_2 molecule. As shown in Figure **9.7.1*b***, this is indeed the case. All the homonuclear alkaline earth diatomic molecules have four valence electrons, which fill both the σ_{ns} bonding orbital and the σ_{ns}* antibonding orbital and give a bond order of 0. Thus Be_2, Mg_2, Ca_2, Sr_2, and Ba_2 are all expected to be unstable, in agreement with experimental data.In the solid state, however, all the alkali metals and the alkaline earth metals exist as extended lattices held together by metallic bonding. At low temperatures, Be_2 is stable.

✔ Example 9.7.1 : Sodium Dimer Ion

Use a qualitative molecular orbital energy-level diagram to predict the valence electron configuration, bond order, and likely existence of the Na_2^- ion.

Given: chemical species

Asked for: molecular orbital energy-level diagram, valence electron configuration, bond order, and stability

Strategy

A. Combine the two sodium valence atomic orbitals to produce bonding and antibonding molecular orbitals. Draw the molecular orbital energy-level diagram for this system.

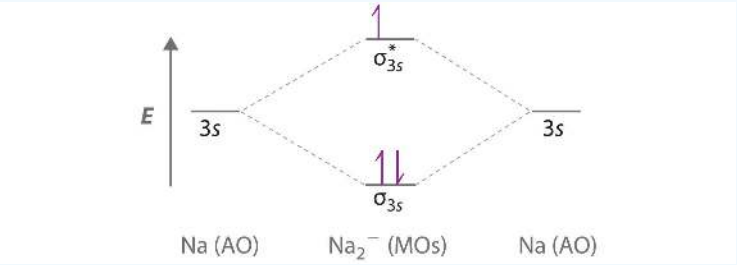
B. Determine the total number of valence electrons in the Na_2^- ion. Fill the molecular orbitals in the energy-level diagram beginning with the orbital with the lowest energy. Be sure to obey the Pauli principle and Hund's rule while doing so.

C. Calculate the bond order and predict whether the species is stable.

Solution

A Because sodium has a [Ne]$3s^1$ electron configuration, the molecular orbital energy-level diagram is qualitatively identical to the diagram for the interaction of two 1*s* atomic orbitals.

B The Na_2^- ion has a total of three valence electrons (one from each Na atom and one for the negative charge), resulting in a filled σ_{3s} molecular orbital, a half-filled σ_{3s*} and a $(\sigma_{3s})^2 (\sigma^*_{3s})^1$ electron configuration.



(CC BY-SA-NC; anonymous).

C The bond order is (2-1)÷2=1/2 With a fractional bond order, we predict that the Na_2^- ion exists but is highly reactive.

? Exercise 9.7.1 : Calcium Dimer Cation

Use a qualitative molecular orbital energy-level diagram to predict the valence electron configuration, bond order, and likely existence of the Ca_2^+ ion.

Molecular Orbitals Formed from n*p* Orbitals

Atomic orbitals other than *ns* orbitals can also interact to form molecular orbitals. Because individual *p*, *d*, and *f* orbitals are not spherically symmetrical, however, we need to define a coordinate system so we know which lobes are interacting in three-dimensional space. Recall that for each *np* subshell, for example, there are np_x, np_y, and np_z orbitals. All have the same energy and are therefore degenerate, but they have different spatial orientations.

$$\sigma_{np_z} = np_z(A) - np_z(B) \tag{9.7.1}$$

Just as with *ns* orbitals, we can form molecular orbitals from *np* orbitals by taking their mathematical sum and difference. When two positive lobes with the appropriate spatial orientation overlap, as illustrated for two np_z atomic orbitals in part (a) in Figure 9.7.2 , it is the mathematical *difference* of their wavefunctions that results in *constructive* interference, which in turn increases the electron probability density between the two atoms. The difference therefore corresponds to a molecular orbital called a σ_{np_z} *bonding molecular orbital* because, just as with the σ orbitals discussed previously, it is symmetrical about the internuclear axis (in this case, the *z*-axis):

$$\sigma_{np_z} = np_z(A) - np_z(B) \tag{9.7.2}$$

The other possible combination of the two np_z orbitals is the mathematical sum:

$$\sigma_{np_z} = np_z(A) + np_z(B) \tag{9.7.3}$$

In this combination, shown in part (b) in Figure 9.7.2 , the positive lobe of one np_z atomic orbital overlaps the negative lobe of the other, leading to *destructive* interference of the two waves and creating a node between the two atoms. Hence this is an antibonding molecular orbital. Because it, too, is symmetrical about the internuclear axis, this molecular orbital is called a $\sigma_{np_z} = np_z(A) - np_z(B)$ *antibonding molecular orbital*. Whenever orbitals combine, *the bonding combination is always lower in energy* (more stable) than the atomic orbitals from which it was derived, and *the antibonding combination is higher in energy* (less stable).

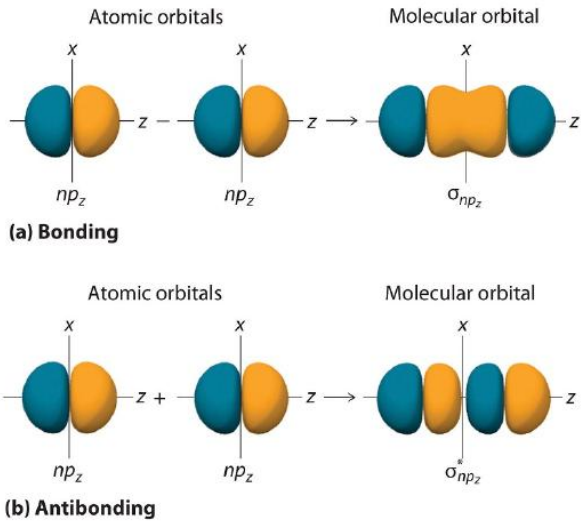


Figure 9.7.2 : Formation of Molecular Orbitals from np_z Atomic Orbitals on Adjacent Atoms.(a) By convention, in a linear molecule or ion, the z-axis always corresponds to the internuclear axis, with +z to the right. As a result, the signs of the lobes of the np_z atomic orbitals on the two atoms alternate − + − +, from left to right. In this case, the σ (bonding) molecular orbital corresponds to the mathematical *difference*, in which the overlap of lobes with the same sign results in increased probability density between the nuclei. (b) In contrast, the σ* (antibonding) molecular orbital corresponds to the mathematical *sum*, in which the overlap of lobes with opposite signs results in a nodal plane of zero probability density perpendicular to the internuclear axis. (CC BY-SA-NC; anonymous).

The remaining *p* orbitals on each of the two atoms, np_x and np_y, do not point directly toward each other. Instead, they are perpendicular to the internuclear axis. If we arbitrarily label the axes as shown in Figure 9.7.3 , we see that we have two pairs of *np* orbitals: the two np_x orbitals lying in the plane of the page, and two np_y orbitals perpendicular to the plane. Although these two pairs are equivalent in energy, the np_x orbital on one atom can interact with only the np_x orbital on the other, and the np_y orbital on one atom can interact with only the np_y on the other. These interactions are side-to-side rather than the head-to-head interactions characteristic of σ orbitals. Each pair of overlapping atomic orbitals again forms two molecular orbitals: one corresponds to the arithmetic sum of the two atomic orbitals and one to the difference. The sum of these side-to-side interactions increases the electron probability in the region above and below a line connecting the nuclei, so it is a bonding molecular orbital that is called a pi (π) orbital (a bonding molecular orbital formed from the side-to-side interactions of two or more parallel np atomic orbitals). The difference results in the overlap of orbital lobes with opposite signs, which produces a nodal plane perpendicular to the internuclear axis; hence it is an antibonding molecular orbital, called a pi star (π*) orbital An antibonding molecular orbital formed from the difference of the side-to-side interactions of two or more parallel *np* atomic orbitals, creating a nodal plane perpendicular to the internuclear axis..

$$\pi_{np_z} = np_x\,(A) + np_x\,(B) \tag{9.7.4}$$

$$\pi^*_{np_z} = np_x\,(A) - np_x\,(B) \tag{9.7.5}$$

The two np_y orbitals can also combine using side-to-side interactions to produce a bonding π_{np_y} molecular orbital and an antibonding $\pi^*_{np_y}$ molecular orbital. Because the np_x and np_y atomic orbitals interact in the same way (side-to-side) and have the same energy, the π_{np_z} and π_{np_y}molecular orbitals are a degenerate pair, as are the $\pi^*_{np_z}$ and $\pi^*_{np_y}$ molecular orbitals.

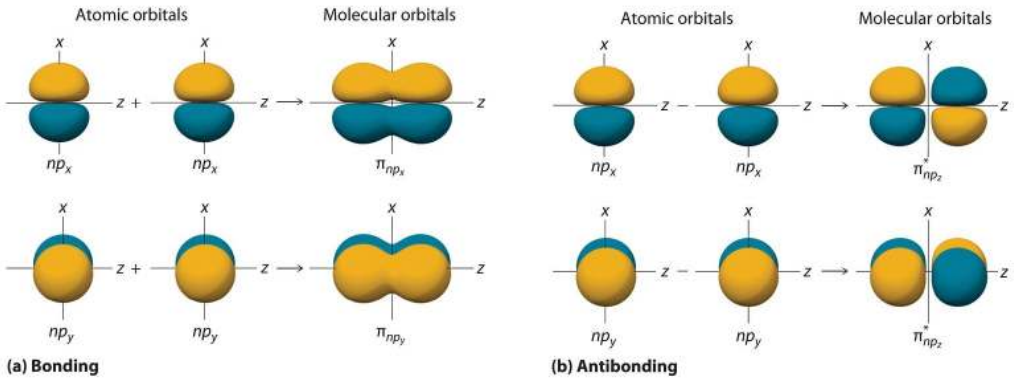


Figure 9.7.3 : Formation of π Molecular Orbitals from npx and npy Atomic Orbitals on Adjacent Atoms.(a) Because the signs of the lobes of both the npx and the npy atomic orbitals on adjacent atoms are the same, in both cases the mathematical sum corresponds to a π (bonding) molecular orbital. (b) In contrast, in both cases, the mathematical difference corresponds to a π* (antibonding) molecular orbital, with a nodal plane of zero probability density perpendicular to the internuclear axis. (CC BY-SA-NC; anonymous).

Energies for Homonuclear Diatomic Molecules

We now describe examples of systems involving period 2 homonuclear diatomic molecules, such as N_2, O_2, and F_2. When we draw a molecular orbital diagram for a molecule, there are four key points to remember:

1. The number of molecular orbitals produced is the same as the number of atomic orbitals used to create them.
2. As the overlap between two atomic orbitals increases, the difference in energy between the resulting bonding and antibonding molecular orbitals increases.
3. When two atomic orbitals combine to form a pair of molecular orbitals, the bonding molecular orbital is stabilized about as much as the antibonding molecular orbital is destabilized.
4. The interaction between atomic orbitals is greatest when they have the same energy.

Figure 9.7.4 is an energy-level diagram that can be applied to two *identical* interacting atoms that have three *np* atomic orbitals each. There are six degenerate *p* atomic orbitals (three from each atom) that combine to form six molecular orbitals, three bonding and three antibonding. The bonding molecular orbitals are lower in energy than the atomic orbitals because of the increased stability associated with the formation of a bond. Conversely, the antibonding molecular orbitals are higher in energy, as shown. The energy difference between the σ and σ* molecular orbitals is significantly greater than the difference between the two π and π* sets. The reason for this is that the atomic orbital overlap and thus the strength of the interaction are greater for a σ bond than a π bond, which means that the σ molecular orbital is more stable (lower in energy) than the π molecular orbitals.

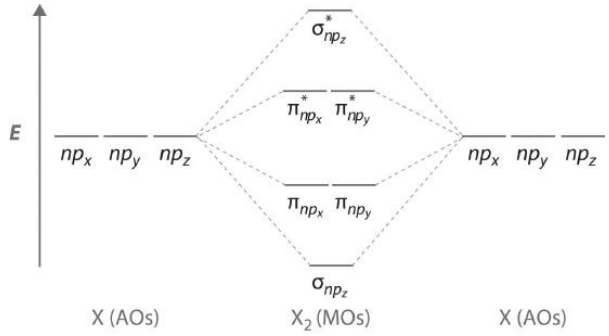


Figure 9.7.4 : The Relative Energies of the σ and π Molecular Orbitals Derived from np_x, np_y, and np_z Orbitals on Identical Adjacent Atoms. Because the two np_z orbitals point directly at each other, their orbital overlap is greater, so the difference in energy between the σ and σ* molecular orbitals is greater than the energy difference between the π and π* orbitals. (CC BY-SA-NC; anonymous).

> *The number of molecular orbitals is always equal to the total number of atomic orbitals we started with.*

We illustrate how to use these points by constructing a molecular orbital energy-level diagram for F_2. We use the diagram in Figure **9.7.5***a*; the $n = 1$ orbitals (σ_{1s} and σ_{1s}*) are located well below those of the $n = 2$ level and are not shown. As illustrated in the diagram, the σ_{2s} and σ_{2s}* molecular orbitals are much lower in energy than the molecular orbitals derived from the 2*p* atomic orbitals because of the large difference in energy between the 2*s* and 2*p* atomic orbitals of fluorine. The lowest-energy molecular orbital derived from the three 2*p* orbitals on each F is σ_{2p_z} and the next most stable are the two degenerate orbitals, π_{2p_x} and π_{2p_y}. For each bonding orbital in the diagram, there is an antibonding orbital, and the antibonding orbital is destabilized by about as much as the corresponding bonding orbital is stabilized. As a result, the $\sigma^*_{2p_z}$ orbital is higher in energy than either of the degenerate $\pi^*_{2p_x}$ and $\pi^*_{2p_y}$ orbitals. We can now fill the orbitals, beginning with the one that is lowest in energy.

Each fluorine has 7 valence electrons, so there are a total of 14 valence electrons in the F_2 molecule. Starting at the lowest energy level, the electrons are placed in the orbitals according to the Pauli principle and Hund's rules. Two electrons each fill the σ_{2s} and σ_{2s}* orbitals, 2 fill the σ_{2p_z} orbital, 4 fill the two degenerate π orbitals, and 4 fill the two degenerate π* orbitals, for a total of 14 electrons.

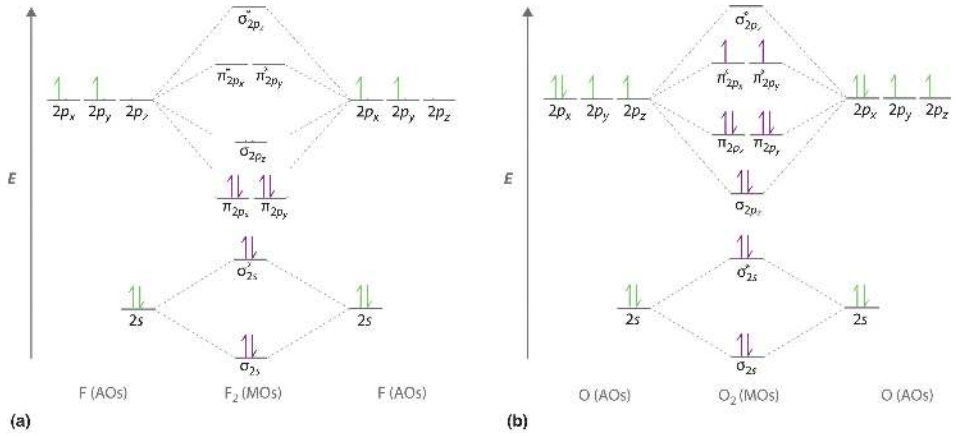


Figure 9.7.5 : Molecular Orbital Energy-Level Diagrams for Homonuclear Diatomic Molecules.(a) For C_2, with 8 valence electrons (4 from each F atom). This diagram shows 6 electrons in bonding orbitals and 2 in antibonding orbitals, resulting in a bond order of 2. (b) For O_2, with 12 valence electrons (6 from each O atom), there are only 2 electrons to place in the $\left(\pi^*_{np_x}, \pi^*_{np_y}\right)$ pair of orbitals. Hund's rule dictates that one electron occupies each orbital, and their spins are parallel, giving the O_2 molecule two unpaired electrons. This diagram shows 8 electrons in bonding orbitals and 4 in antibonding orbitals, resulting in a predicted bond order of 2. (CC BY-SA-NC; anonymous).

For period 2 diatomic molecules to the left of N_2 in the periodic table, a slightly different molecular orbital energy-level diagram is needed because the σ_{2p_z} molecular orbital is slightly *higher* in energy than the degenerate $\pi^*_{np_x}$ and $\pi^*_{np_y}$ orbitals. The difference in energy between the 2*s* and 2*p* atomic orbitals increases from Li_2 to F_2 due to increasing nuclear charge and poor screening of the 2*s* electrons by electrons in the 2*p* subshell. The bonding interaction between the 2*s* orbital on one atom and the $2p_z$ orbital on the other is most important when the two orbitals have similar energies. This interaction decreases the energy of the σ_{2s} orbital and increases the energy of the σ_{2p_z} orbital. Thus for Li_2, Be_2, B_2, C_2, and N_2, the σ_{2p_z} orbital is higher in energy than the σ_{3p_z} orbitals, as shown in Figure 9.7.6 . Experimentally, the energy gap between the *ns* and *np* atomic orbitals *increases* as the nuclear charge increases (Figure 9.7.6). Thus for example, the σ_{2p_z} molecular orbital is at a lower energy than the $\pi_{2p_{x,y}}$ pair.

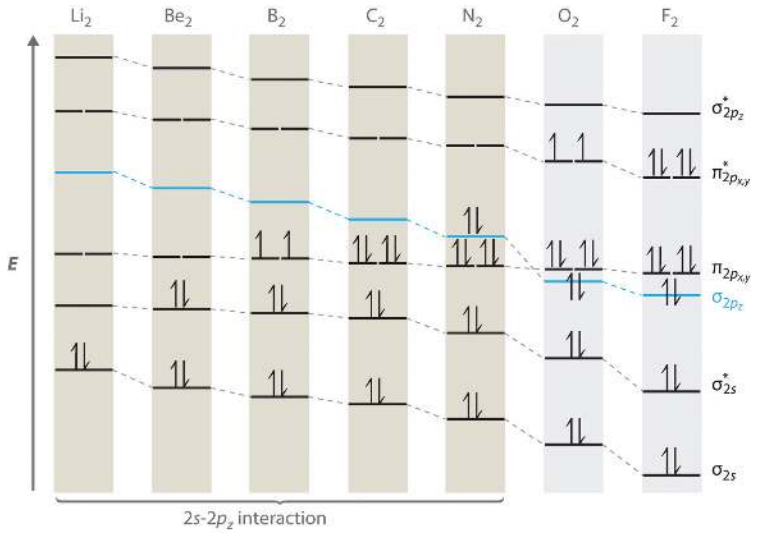


Figure 9.7.6 : Molecular Orbital Energy-Level Diagrams for the Diatomic Molecules of the Period 2 Elements. Unlike earlier diagrams, only the molecular orbital energy levels for the molecules are shown here. For simplicity, the atomic orbital energy levels for the component atoms have been omitted. For Li_2 through N_2, the σ_{2p_z} orbital is higher in energy than the $\pi_{2p_{x,y}}$ orbitals. In contrast, the σ_{2p_z} orbital is lower in energy than the $\pi_{2p_{x,y}}$ orbitals for O_2 and F_2 due to the increase in the energy difference between the 2s and 2*p* atomic orbitals as the nuclear charge increases across the row. (CC BY-SA-NC; anonymous).

Example 9.7.2 : Diatomic Sulfur

Use a qualitative molecular orbital energy-level diagram to predict the electron configuration, the bond order, and the number of unpaired electrons in S_2, a bright blue gas at high temperatures.

Given: chemical species

Asked for: molecular orbital energy-level diagram, bond order, and number of unpaired electrons

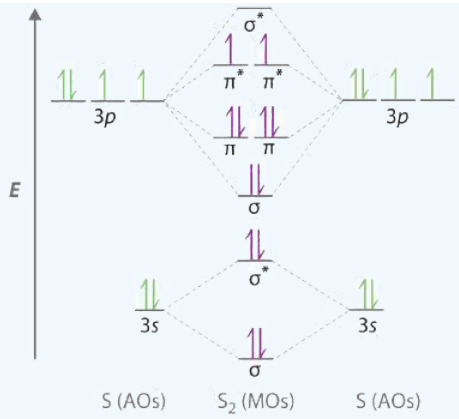
Strategy:

A. Write the valence electron configuration of sulfur and determine the type of molecular orbitals formed in S_2. Predict the relative energies of the molecular orbitals based on how close in energy the valence atomic orbitals are to one another.
B. Draw the molecular orbital energy-level diagram for this system and determine the total number of valence electrons in S_2.
C. Fill the molecular orbitals in order of increasing energy, being sure to obey the Pauli principle and Hund's rule.
D. Calculate the bond order and describe the bonding.

Solution:

A Sulfur has a $[Ne]3s^23p^4$ valence electron configuration. To create a molecular orbital energy-level diagram similar to those in Figures 9.7.6 and 9.7.7 , we need to know how close in energy the 3*s* and 3*p* atomic orbitals are because their energy separation will determine whether the $\pi_{3p_{x,y}}$ or the σ_{3p_z}> molecular orbital is higher in energy. Because the *ns*–*np* energy gap *increases* as the nuclear charge increases, the σ_{3p_z} molecular orbital will be lower in energy than the $\pi_{3p_{x,y}}$ pair.

B The molecular orbital energy-level diagram is as follows:



(CC BY-SA-NC; anonymous).

Each sulfur atom contributes 6 valence electrons, for a total of 12 valence electrons.

C Ten valence electrons are used to fill the orbitals through π_{3p_z} and π_{3p_y}, leaving 2 electrons to occupy the degenerate $\pi^*_{3p_z}$ and $\pi^*_{3p_y}$ pair. From Hund's rule, the remaining 2 electrons must occupy these orbitals separately with their spins aligned. With the numbers of electrons written as superscripts, the electron configuration of S_2 is $(\sigma_{3s})^2(\sigma^*_{3s})^2(\sigma_{3p_z})^2(\pi_{3p_{x,y}})^4(\pi^*_{3p_{x,y}})^2$ with 2 unpaired electrons. The bond order is $(8 - 4) \div 2 = 2$, so we predict an S=S double bond.

? Exercise 9.7.2

Use a qualitative molecular orbital energy-level diagram to predict the electron configuration, the bond order, and the number of unpaired electrons in the peroxide ion (O_2^{2-}).

Molecular Orbitals Formed from *ns* with n*p* Orbitals

Although many combinations of atomic orbitals form molecular orbitals, we will discuss only one other interaction: an *ns* atomic orbital on one atom with an np_z atomic orbital on another. As shown in Figure 9.7.7 , the sum of the two atomic wavefunctions ($ns + np_z$) produces a σ bonding molecular orbital. Their difference ($ns - np_z$) produces a σ* antibonding molecular orbital, which has a nodal plane of zero probability density perpendicular to the internuclear axis.

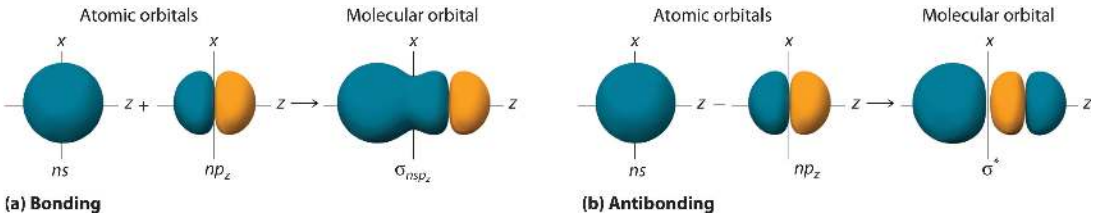


Figure 9.7.7 : Formation of Molecular Orbitals from an ns Atomic Orbital on One Atom and an np_z Atomic Orbital on an Adjacent Atom.(a) The mathematical sum results in a σ (bonding) molecular orbital, with increased probability density between the nuclei. (b) The mathematical difference results in a σ* (antibonding) molecular orbital, with a nodal plane of zero probability density perpendicular to the internuclear axis. (CC BY-SA-NC; anonymous).

9.7: Molecular Orbitals Can Be Ordered According to Their Energies is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **9.7: Molecular Orbitals Can Be Ordered According to Their Energies** is licensed CC BY-NC-SA 4.0.

9.8: Molecular-Orbital Theory Does not Predict a Stable Diatomic Helium Molecule

Learning Objectives

- Using bond order as a metric for the existance of molecules

Bond Order

In the Lewis electron structures, the number of electron pairs holding two atoms together was called the *bond order*. Within the molecular orbital approach, *bond order* is defined as one-half the *net* number of bonding electrons:

$$\text{bond order} = \frac{\text{number of bonding electrons} - \text{number of antibonding electrons}}{2} \tag{9.8.1}$$

To calculate the bond order of H_2, we know that the σ_{1s} (bonding) molecular orbital contains two electrons, while the σ^*_{1s} (antibonding) molecular orbital is empty. The bond order of H_2 is therefore

$$\text{bond order} = \frac{2-0}{2} = 1 \tag{9.8.2}$$

This result corresponds to the single covalent bond; double and triple bonds contain four or six electrons, respectively, and correspond to bond orders of 2 and 3.

We can use energy-level diagrams to describe the bonding in other pairs of atoms and ions where n = 1, such as the H_2^+ ion, the He_2^+ ion, and the He_2 molecule. Again, we fill the lowest-energy molecular orbitals first while being sure not to violate the Pauli principle or Hund's Rules.

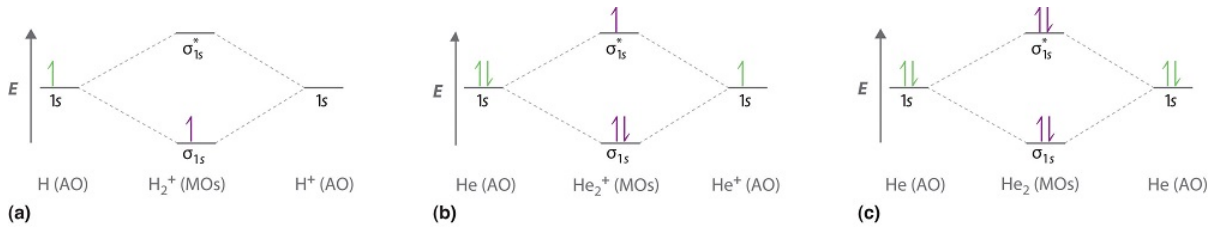


Figure 9.8.1 : Molecular Orbital Energy-Level Diagrams for Diatomic Molecules with Only 1s Atomic Orbitals. (a) The H_2^+ ion, (b) the He_2^+ ion, and (c) the He_2 molecule are shown here. (CC BY-SA-NC; anonymous)

Figure **9.8.1***a* shows the energy-level diagram for the H_2^+ ion, which contains two protons and only one electron. The single electron occupies the σ_{1s} bonding molecular orbital, giving a $(\sigma_{1s})^1$ electron configuration. The number of electrons in an orbital is indicated by a superscript. In this case, the bond order is (via Equation 9.8.1)

$$\text{bond order} = \frac{1-0}{2} = 1/2$$

Because the bond order is greater than zero, the H_2^+ ion should be more stable than an isolated H atom and a proton. We can therefore use a molecular orbital energy-level diagram and the calculated bond order to predict the relative stability of species such as H_2^+. With a bond order of only 1/2 the bond in H_2^+ should be weaker than in the H_2 molecule, and the H–H bond should be longer. As shown in Table 9.8.1 , these predictions agree with the experimental data.

Table 9.8.1 : Molecular Orbital Electron Configurations, Bond Orders, Bond Lengths, and Bond Energies for some Simple Homonuclear Diatomic Molecules and Ions

Species	Electron Configuration	Bond Order	Bond Length (pm)	Bond Energy (kJ/mol)
H_2^+	$(\sigma_{1s})^1$	1/2	106	269
H_2	$(\sigma_{1s})^2$	1	74	436
He_2^+	$(\sigma_{1s})^2(\sigma^*_{1s})^1$	1/2	108	251
He_2	$(\sigma_{1s})^2(\sigma^*_{1s})^2$	0	5,500	4.6×10^{-5}

Figure 9.8.1*b* is the molecular orbital energy-level diagram for He_2^{2+}. This ion has a total of three valence electrons. Because the first two electrons completely fill the σ_{1s} molecular orbital, the Pauli principle states that the third electron must be in the σ_{1s}^* antibonding orbital, giving a $(\sigma_{1s})^2(\sigma_{1s}^*)^1$ electron configuration. This electron configuration gives a bond order (via Equation 9.8.1) of

$$\text{bond order} = \frac{2-1}{2} = 1/2$$

As with H_2^+, the He_2^+ ion should be stable, but the He–He bond should be weaker and longer than in H_2. In fact, the He_2^+ ion can be prepared, and its properties are consistent with our predictions (Table 9.8.1).

He_2^{2+} ✔ Example 9.8.1 : The ion

Use a molecular orbital energy-level diagrams to predict the bond order and stability of the He_2^{2+} ion.

Given: chemical species

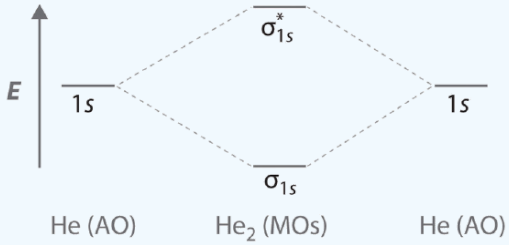
Asked for: molecular orbital energy-level diagram, bond order, and stability

Strategy:

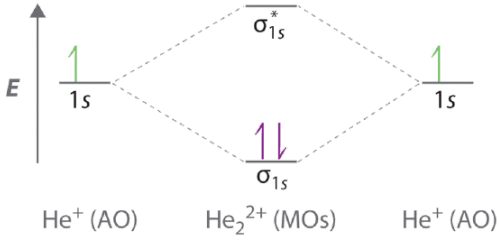
A. Combine the two He valence atomic orbitals to produce bonding and antibonding molecular orbitals. Draw the molecular orbital energy-level diagram for the system.
B. Determine the total number of valence electrons in the He_2^{2+} ion. Fill the molecular orbitals in the energy-level diagram beginning with the orbital with the lowest energy. Be sure to obey the Pauli principle and Hund's rule while doing so.
C. Calculate the bond order and predict whether the species is stable.

Solution:

A Two He 1*s* atomic orbitals combine to give two molecular orbitals: a σ_{1s} bonding orbital at lower energy than the atomic orbitals and a σ_{1s}^* antibonding orbital at higher energy. The bonding in any diatomic molecule with two He atoms can be described using the following molecular orbital diagram:



B The He_2^{2+} ion has only two valence electrons (two from each He atom minus two for the +2 charge). We can also view He_2^{2+} as being formed from two He^+ ions, each of which has a single valence electron in the 1*s* atomic orbital. We can now fill the molecular orbital diagram:



The two electrons occupy the lowest-energy molecular orbital, which is the bonding (σ_{1s}) orbital, giving a $(\sigma_{1s})^2$ electron configuration. To avoid violating the Pauli principle, the electron spins must be paired.

C So the bond order is (via Equation 9.8.1)

$$\frac{2-0}{2}=1$$

He_2^{2+} is therefore predicted to contain a single He–He bond. Thus it should be a stable species.

H_2^{2-}? Exercise 9.8.1 : The Ion

Use a molecular orbital energy-level diagram to predict the valence-electron configuration and bond order of the H_2^{2-} ion. Is this a stable species?

The Helium Dimer

Finally, we examine the He_2 molecule, formed from two He atoms with $1s^2$ electron configurations. Figure **9.8.1c** is the molecular orbital energy-level diagram for He_2. With a total of four valence electrons, both the σ_{1s} bonding and σ^*_{1s} antibonding orbitals must contain two electrons. This gives a $(\sigma_{1s})^2\,(\sigma^*_{1s})^1$ electron configuration, with a predicted bond order (via Equation 9.8.1) of

$$\text{bond order}=\frac{2-2}{2}=0 \tag{9.8.3}$$

which indicates that the He_2 molecule has no net covalent bond and is not a stable species.

The ability to track bond order to bond strength is due to the fact that the energy difference between the anti-bonding σ^*_{1s} molecular orbital and the original $1s$ atomic orbital is *larger* than the energy difference between the bonding σ_{1s} molecular orbital and the $1s$ atomic orbitals. This was derived previously where the stabilization energy (ΔE_+) of the σ_{1s} molecular orbital is less than the destabilization energy (ΔE_-) of the anti-bonding σ^*_{1s} molecular orbital:

$$\Delta E_\pm = E_\pm - E_H \tag{9.8.4}$$

$$= \frac{e^2}{4\pi\epsilon_0 R}+\frac{J\pm K}{1\pm S} \tag{9.8.5}$$

Hence, the anti-bonding σ^*_{1s} molecular orbital is more destabilized relative to the atomic orbitals than the bonding σ_{1s} molecular orbital is stabilized relative to the $1s$ atomic orbitals (E_H.

The fact that the anti-bonding MO energy difference is larger than the bonding σ_{1s} molecular orbital energy difference is the true reason that helium dimer is not predicted to exist with a covalent bond. Of the four valence electrons in helium dimer, two will fill the bonding σ_{1s} molecular orbital, and the other two will fill the anti-bonding σ^*_{1s} molecular orbital. The two electrons in the bonding σ_{1s} molecular orbital will achieve some stabilization relative to the $1s$ atomic orbitals, but the two electrons in the anti-bonding σ^*_{1s} molecular orbital will achieve *greater* de-stabilization relative to their position in the atomic orbitals. The net result is a less stable molecule than if the electrons remained in their respective $1s$ atomic orbitals.

> *The electrons in antibonding orbitals cancel (and exceed) the stabilization resulting from electrons in bonding orbitals. Consequently, any system that has equal numbers of bonding and antibonding electrons will have a bond order of 0, and it is predicted to be unstable and therefore not to exist in nature (at least as a covalently bonding complex).*

However, Van der Waals Helium Dimers do exist

Based on molecular orbital theory discussed above, the He_2 molecule should not exist since no covalent bond formed between the helium atoms (Equation 9.8.3). However, the molecular orbital description above neglects the van der Waals force that exists between the atoms as demonstrated by the existence of liquid helium (at 4 K). So a "molecule" composed of two helium atoms bound by the van der Waals force may exist by this attractive force instead - and it does.

A helium dimer molecule bound by Van der Waals forces was first proposed by John Clarke Slater in 1928 and observed in 1993 by Gentry and coworkers. Interestingly, He_2 is the largest known molecule of two atoms when in its ground state with an extremely long bond length with a separation of about 5,200 pm. The binding energy is only $4.6\times10^{-5}\ kJ/mol$, so the **He—He** bond is 5,000 times weaker than the covalent bond in the hydrogen molecule (Table 9.8.1).

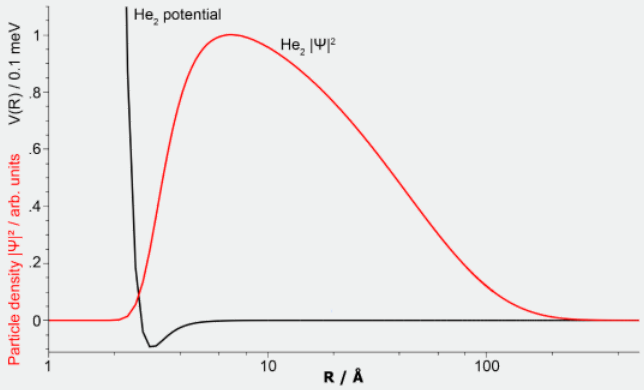


Figure 9.8.2 : The weak helium-helium Van der Waals potential decrease leads to the particle probability density distribution leaking more into the classically forbidden region (i..e, tunneling). This effect allows the wavefunction to extend to sizes of fullerenes, the diameter of DNA and even small viruses ($\mathbf{He_2}$ potential and wavefunction taken from Przybytek 2010), while the classical turning point is located at 13.6 Å, the overall wavefunction extends to more than 200 Å. (Inspire; http://inspirehep.net/record/1415151/plots)

Conclusion

The decrease in energy caused by the bonding orbital (constructive interference of the atomic orbitals) is canceled by the increase in energy caused by the antibonding orbital (destructive interference of the atomic orbitals), so it is not energetically favorable for the helium atoms to be in such proximity, so if that situation arises, they'll separate quickly since there's no force keeping them there.

9.8: Molecular-Orbital Theory Does not Predict a Stable Diatomic Helium Molecule is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **9.8: Molecular-Orbital Theory Does not Predict a Stable Diatomic Helium Molecule** is licensed CC BY-NC-SA 4.0.

9.9: Electrons Populate Molecular Orbitals According to the Pauli Exclusion Principle

Just like with multielectron atoms, there are restrictions on how electrons occupy molecular orbitals. The most important of these rules is the Pauli exclusion principle, which we will not try to prove but will take as a fact. In general chemistry, you were told that the Pauli exclusion principle means that electrons in the same orbital must have opposite spin, and that only two electrons can be in each orbital. That is close, but not entirely accurate. The true statement of the Pauli exclusion principle is this: wavefunctions must be antisymmetric with respect to exchange of identical fermions, and symmetric with respect to exchange of identical bosons. A is a particle with half-integer spin, and a is a particle with integer spin. Recall that electrons are fermions (spin 1/2): we represent the two possible spin states as $|\alpha\rangle$ ($m_z = 1/2$) and $|\beta\rangle$ ($m_z = -1/2$), according to their spin projections.

Because all electrons are identical, the Pauli exclusion principle requires that the electronic wavefunction be antisymmetric with respect to exchanging any of them.[2] The word "antisymmetric" in this context means that if we swap two electrons (i.e., exchange their labels), the wavefunction must change sign. To denote both the orbital of an electron and its spin, we use a product of , each of the form $|\psi\sigma\rangle$, where ψ depends only on the spatial coordinates of the electrons, and σ depends only on their spins.

In the ground electronic state of , the electrons both occupy the $|1\sigma\rangle$ spatial orbital. Naively, we could write the following product of spin orbitals:

$$|\psi\rangle = |1\sigma^1\alpha^1\rangle|1\sigma^2\beta^2\rangle \tag{9.9.1}$$

where the superscript refers to the electron number. Note that since the spatial and spin coordinates are independent of one another, we can group them into kets in whichever way is most convenient for us at the time. Let us define the permutation operator $\hat{p}_{12}$, which swaps the labels of electrons 1 and 2, in order to test the wavefunction for asymmetry in accordance with the **Pauli Exclusion Principle**. It is clear that the electronic wavefunction we wrote above does not change sign under this permutation operation!

$$\hat{p}_{12}|\psi\rangle = |1\sigma^2\alpha^2\rangle|1\sigma^1\beta^1\rangle \neq -|\psi\rangle. \tag{9.9.2}$$

Instead, we can take a linear combination of spin functions to make a new wavefunction:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|1\sigma^1\alpha^1 1\sigma^2\beta^2\rangle - |1\sigma^2\alpha^2 1\sigma^1\beta^1\rangle\right) \tag{9.9.3}$$

Since the Hamiltonian did not contain any terms that depend on spin, the energy of this new wavefunction is equal to the energy of our original wavefunction.[3] We can use the permutation operator $\hat{p}_{12}$ to see that this wavefunction is in fact antisymmetric:

$$\begin{aligned}\hat{p}_{12}|\psi\rangle &= \frac{1}{\sqrt{2}}\left(|1\sigma^2\alpha^2 1\sigma^1\beta^1\rangle - |1\sigma^1\alpha^1 1\sigma^2\beta^2\rangle\right)\\ &= -\frac{1}{\sqrt{2}}\left(|1\sigma^1\alpha^1 1\sigma^2\beta^2\rangle - |1\sigma^2\alpha^2 1\sigma^1\beta^1\rangle\right)\\ &= -|\psi\rangle.\end{aligned} \tag{9.9.4}$$

We see here that the orbital is not occupied by one electron of each spin, but in fact it is occupied electrons in a superposition of spin states. In larger systems, antisymmetrized wavefunctions show that each orbital is occupied by a superposition of all electrons in the system in both spin states. Although that is difficult to conceptualize, in most practical cases the electrons can still be thought of as localized in one particular orbital with one particular spin.

There is a standard procedure to generate an appropriately antisymmetrized electronic wavefunction from any set of single-electron functions. Like in multi-electron atoms, the general method for doing this is by use of a Slater determinant: a determinant in which each column contains the product of one molecular orbital and spin function for each possible labeling of the electrons. For our system above (two electrons in $|1\sigma\rangle$, one with spin α, and the other with spin β), we would write

$$|\psi\rangle = \frac{1}{\sqrt{2!}}\begin{vmatrix} |1\sigma^1\alpha^1\rangle & |1\sigma^1\beta^1\rangle \\ |1\sigma^2\alpha^2\rangle & |1\sigma^2\beta^2\rangle \end{vmatrix}, \tag{9.9.5}$$

where once again the superscript indicated the label on the electron. The Slater determinant therefore contains all of the possible ways that we could assign numbers to the electrons in the system.

Note

A few things to know about Slater determinants:

- If two columns are swapped, the determinant changes sign (this is why it satisfies the exclusion principle!).
- If any two columns in a determinant are the same, the determinant is 0. This is why the Pauli exclusion principle only allows one electron of each spin to occupy a single orbital, which you learned way back in Gen Chem. It also directly leads to the "Aufbau" principle: when populating molecular orbitals, we fill up orbitals in order of increasing energy with two electrons each.
- The energy of a Slater determinant (in the independent particle approximation) is simply the sum of the orbitals that make it up (i.e., add up the energy of each electron in the MO diagram).
- The wavefunction described by a Slater determinant is normalized if the individual orbitals that make it up are normalized.
- Two Slater determinants are orthogonal if they differ in any single orbital.

Given an MO diagram, it is easy to generate a Slater determinant for the ground electronic state of a closed-shell molecule using general chemistry rules: the Aufbau principle and the Pauli exclusion principle. Each orbital can accommodate one electrons of each spin, and electrons are placed into orbitals from lowest to highest energy. A system with N electrons in the orbitals $|\phi_1\rangle$, $|\phi_2\rangle$, etc., will generate an $N \times N$ dimensional Slater determinant, and the general form is (if N is even):

$$|\psi\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} |\phi_1^1\alpha^1\rangle & |\phi_1^1\beta^1\rangle & |\phi_2^1\alpha^1\rangle & |\phi_2^1\beta^1\rangle & \cdots & |\phi_{N/2}^1\alpha^1\rangle & |\phi_{N/2}^1\beta^1\rangle \\ |\phi_1^2\alpha^2\rangle & |\phi_1^2\beta^2\rangle & |\phi_2^2\alpha^2\rangle & |\phi_2^2\beta^2\rangle & \cdots & |\phi_{N/2}^2\alpha^2\rangle & |\phi_{N/2}^2\beta^2\rangle \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ |\phi_1^N\alpha^N\rangle & |\phi_1^N\beta^N\rangle & |\phi_2^N\alpha^N\rangle & |\phi_2^N\beta^N\rangle & \cdots & |\phi_{N/2}^N\alpha^N\rangle & |\phi_{N/2}^N\beta^N\rangle \end{vmatrix} \tag{9.9.6}$$

Once again, each column contains the orbital and spin state of one of the electrons in the diagram, and each row in the column corresponds to a different choice of label (1, 2, 3, etc...). Evaluating the determinant yields a normalized, antisymmetrized wavefunction. Therefore, an MO diagram can stand as an effective substitute for an electronic wavefunction; it is implied that the true wavefunction used in calculations comes from the Slater determinant described by the diagram.

9.9: Electrons Populate Molecular Orbitals According to the Pauli Exclusion Principle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.10: Molecular Orbital Theory Predicts that Molecular Oxygen is Paramagnetic

Learning Objectives

- To describe the connection between bond order, bond length and bond energy in diatomic molecules
- To explain the observed paramagnetic properties of molecular oxygen with Molecular Orbital theory

In general chemistry courses, students learn that covalent bonds can come as either single, double or triple bonds, which are identifies by their bond order. Both bond length and bond energy changes as the bond order increases and as the number of electrons shared between two atoms in a molecule increases, the bond order of a bond increases, the strength of the bond increases and the distance between nuclei decreases (Table 9.10.1).

Table 9.10.1 : General Correlation between Bond Strength, length and order in Covalent bonds

Bond	Bond Order	Bond Enthalpy (kJ/mol)	Bond Length (Å)
C–C	1	348	1.54
C=C	2	614	1.34
C≡C	3	839	1.20
N–N	1	163	1.47
N=N	2	418	1.24
N≡N	3	941	1.10

The above trend can be observed in the first row diatomics in Figure 9.10.1 . The bond order can be determined directly form the molecular orbital electron configurations. For diatomics, the occupations can correlate to bond length, bond energies (Figure 9.10.1).

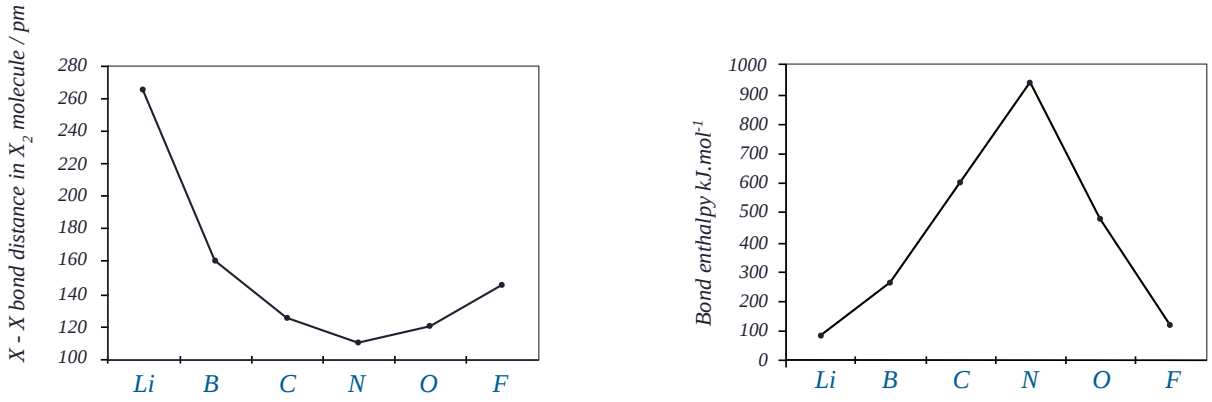


Figure 9.10.1 : Plot of bond length (left) and bond energy (right) for first row diatomics. (CC BY-NC; Ümit Kaya via LibreTexts)

The trends in Figure 9.10.1 and Table 9.10.1 extend to molecular ions.

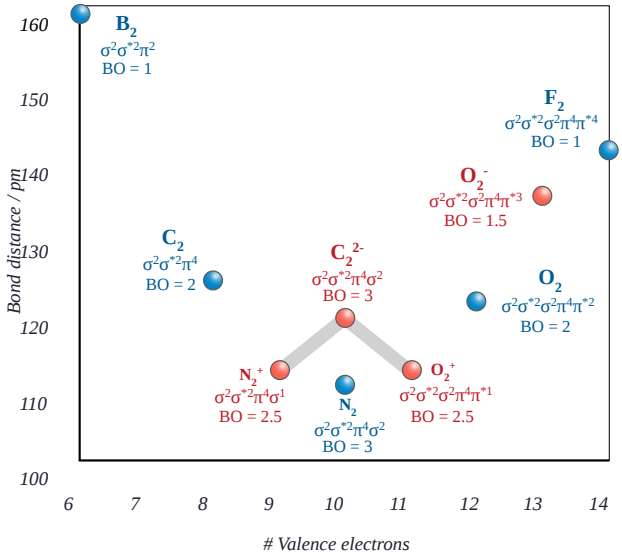


Figure 9.10.2 : Plot of bond length (left) and bond energy (right) for first row diatomics. (CC BY-NC; Ümit Kaya via LibreTexts)

Example 9.10.1 : Molecular Oxygen

Arrange the following four molecular oxygen species in order of increasing bond length: O_2^+, O_2, O_2^-, and O_2^{2-}.

Solution

The bond length in the oxygen species can be explained by the positions of the electrons in molecular orbital theory. To obtain the molecular orbital energy-level diagram for O_2, we need to place 12 valence electrons (6 from each O atom) in the energy-level diagram shown in Figure 9.10.1 . We again fill the orbitals according to Hund's rules and the Pauli principle, beginning with the orbital that is lowest in energy. Two electrons each are needed to fill the σ_{2s} and σ_{2s}* orbitals, two more to fill the σ_{2p_z} orbital, and 4 to fill the degenerate $\pi^*_{2p_x}$ and $\pi^*_{2p_y}$ orbitals. According to Hund's first rule, the last 2 electrons must be placed in separate π^* orbitals with their spins parallel, giving a multiplicity of 3 (a triplet state) with two unpaired electrons. This leads to a predicted bond order of

$$\frac{8-4}{2}=2$$

which corresponds to a double bond, in agreement with experimental data: the O–O bond length is 120.7 pm, and the bond energy is 498.4 kJ/mol at 298 K.

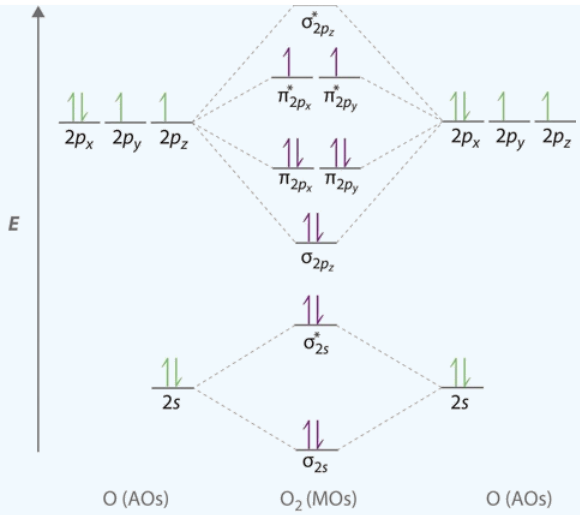


Figure 9.10.3 : Molecular Orbital Energy-Level Diagrams for O_2. With 12 valence electrons (6 from each O atom), there are only 2 electrons to place in the $\left(\pi^*_{np_z}, \pi^*_{np_y}\right)$ pair of orbitals. Hund's first rule dictates that one electron occupies each orbital, and their spins are parallel, giving the O_2 molecule two unpaired electrons. This diagram shows 8 electrons in bonding orbitals and 4 in antibonding orbitals, resulting in a predicted bond order of 2. (CC BY-SA-NC; Anonymous by request).

The bond order is determined from the the electron configurations. The electron configurations for the four species are contrasted below.

- O_2:

$$\sigma_{1s}^2\sigma_{1s}^{*\,2}\sigma_{2s}^2\sigma_{2s}^{*\,2}\sigma_{2p}^2\pi_{2p_y}^2\pi_{2p_y}^{*\,2}\pi_{2p_z}^1\pi_{2p_y}^{*\,1}$$

From Equation ???, the bond order for O_2 is 2 (i.e., a double bond).

- O_2^+:

$$\sigma_{1s}^2\sigma_{1s}^{*\,2}\sigma_{2s}^2\sigma_{2s}^{*\,2}\sigma_{2p}^2\pi_{2p_y}^2\pi_{2p_y}^{*\,2}\pi_{2p_z}^1\pi_{2p_y}^{*\,0}$$

From Equation ???, the bond order for O_2^+ is 2.5. An alternative and equally valid configuration is

$$\sigma_{1s}^2\sigma_{1s}^{*\,2}\sigma_{2s}^2\sigma_{2s}^{*\,2}\sigma_{2p}^2\pi_{2p_y}^2\pi_{2p_y}^{*\,2}\pi_{2p_z}^0\pi_{2p_y}^{*\,1}$$

- O_2^-:

$$\sigma_{1s}^2\sigma_{1s}^{*\,2}\sigma_{2s}^2\sigma_{2s}^{*\,2}\sigma_{2p}^2\pi_{2p_y}^2\pi_{2p_y}^{*\,2}\pi_{2p_z}^2\pi_{2p_y}^{*\,1}$$

From Equation ???, the bond order for O_2^- is 1.5. An alternative and equally valid configuration is

$$\sigma_{1s}^2\sigma_{1s}^{*\,2}\sigma_{2s}^2\sigma_{2s}^{*\,2}\sigma_{2p}^2\pi_{2p_y}^2\pi_{2p_y}^{*\,2}\pi_{2p_z}^1\pi_{2p_y}^{*\,2}$$

- O_2^{2-}:

$$\sigma_{1s}^2\sigma_{1s}^{*\,2}\sigma_{2s}^2\sigma_{2s}^{*\,2}\sigma_{2p}^2\pi_{2p_y}^2\pi_{2p_y}^{*\,2}\pi_{2p_z}^2\pi_{2p_y}^{*\,2}$$

From Equation ???, the bond order for O_2^{2-} is 1.

The bond order decreases and the bond length increases in the order. The predicted order of increasing bondlength then is $O_2^+ < O_2 < O_2^- < O_2^{2-}$. This trend is confirmed experimentally with O_2^+ (112.2 pm), O_2 (121 pm), O_2^- (128 pm) and O_2^{2-} (149 pm).

Molecular Oxygen is Paramagnetic

We now turn to a molecular orbital description of the bonding in O_2. It so happens that the molecular orbital description of this molecule provided an explanation for a long-standing puzzle that could not be explained using other bonding models. None of the other bonding models (e.g., Valence Bond theory or Lewis bonding) can predict the presence of two unpaired electrons in O_2.

Chemists had long wondered why, unlike most other substances, liquid O_2 is attracted into a magnetic field. As shown in Video 9.10.1 , it actually remains suspended between the poles of a magnet until the liquid boils away. The only way to explain this behavior was for O_2 to have unpaired electrons, making it paramagnetic. This result was one of the earliest triumphs of molecular orbital theory over the other bonding approaches.



Video 9.10.1 : Liquid O_2 Suspended between the Poles of a Magnet. Because the O_2 molecule has two unpaired electrons, it is paramagnetic. Consequently, it is attracted into a magnetic field, which allows it to remain suspended between the poles of a powerful magnet until it evaporates.

Advanced: Spin Barriers

The magnetic properties of O_2 are not just a laboratory curiosity; they are absolutely crucial to the existence of life. Because Earth's atmosphere contains 20% oxygen, all organic compounds, including those that compose our body tissues, should react rapidly with air to form H_2O, CO_2, and N_2 in an exothermic reaction. Fortunately for us, however, this reaction is very, very slow. The reason for the unexpected stability of organic compounds in an oxygen atmosphere is that virtually all organic compounds, as well as H_2O, CO_2, and N_2, have only paired electrons, whereas oxygen has two unpaired electrons. Thus the reaction of O_2 with organic compounds to give H_2O, CO_2, and N_2 would require that at least one of the electrons on O_2 change its spin during the reaction. This would require a large input of energy, an obstacle that chemists call a *spin barrier*.

9.10: Molecular Orbital Theory Predicts that Molecular Oxygen is Paramagnetic is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

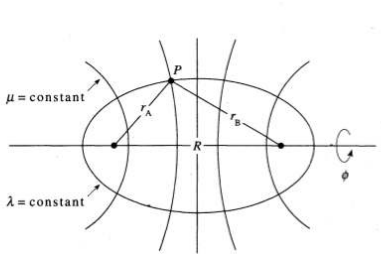
- **9.10: Molecular Orbital Theory Predicts that Molecular Oxygen is Paramagnetic** is licensed CC BY-NC-SA 4.0.

9.E: The Chemical Bond: Diatomic Molecules (Exercises)

Solutions to select questions can be found online.

9.3

The overlap integral and other integrals that arise in two-center systems like that of H_2 which is called a **two-center integral**. The two center integrals are much easier being evaluated by simply using the coordinate system, elliptic coordinates. In the coordinate system, where there are simply two fixed points which are separated by distance R.



A point P is given by three coordinates which are

- $\lambda = r_A + r_B/R$
- $\mu = r_A + r_B/R$

the angle ø is the angle that (r_A, r_B, R) triangle makes the interfocal axis. The differential volume element (elliptic coordinates) $\Leftrightarrow$ $dr = R^3/8(\lambda^2 - \mu^2)d\lambda\, d\mu\, d\phi$

Given definitions λ, μ, ϕ → prove that

- $1 \leq \lambda < \infty$
- $-1 \leq \mu \leq 1$
- $0 \leq \phi \leq 2\pi$

Use elliptic coordinates to evaluate the overlap integral.

Use Equation:

$$S = \int dr\, 1s_A 1s_B = \frac{Z^3}{\pi} \int dr\, e^{-Zr_A} e^{-Zr_B}$$

9.4

Repeat the calculation in 9.3 for $\psi_- = 1s_A - 1s_B$

9.9

Show that the $2p_x$ and the $2p_y$ orbitals do not overlap. Use $2p_x \equiv C\cos\theta$ and $2p_y \equiv C\sin\theta\cos\theta$ and $f(r)$ as the radial component.

9.10

what does the overlap integral represent?

9.14

Which of the three species is the is the least stable due to bond order: O_2, O_2^+, O_2^{2-}. Hint it may be helpful to draw molecular orbitals for each species although it is not required.

9.17

Write the electron configuration and bond order of carbon monoxide.

9.17

Determine the bond order diphosphorus P_2.

9.18

Determine the bond order for the bond in nitrogen oxide cation, NO^+.

9.18

What is the electron configuration and bond order of nitric oxide ion?

9.21

Determine the bond order for the NO molecule.

9.22

How is the energy-level diagram for diatomic OH radical differ from that of NH? What is the highest occupied molecular orbital of OH?

9.23

If a light source generates a light at 73.4 nm while a photoelectron spectrometer is measuring the kinetic energy of the electrons ionized when the molecule absorbs this light, what is the largest electron binding energy that can be measured? Also, is it possible to determine the energy of the occupied molecular orbitals of a molecule using the kinetic energy of the ionized electrons measured? If yes, explain. *Hint* : Recall the photoelectron effect discussed in lecture 2.

9.24

Why would the ionization energy of a $3p$ electron from HCl molecule be lower than a $3p$ electron from a chlorine atom?

9.27

Photoelectron spectroscopy involves the measurement of kinetic energy of photoelectrons to determine the binding energy, intensity and angular distributions of these electrons and use the information obtained to examine the electronic structure of molecules. It differs from the conventional methods of spectroscopy in that it detects electrons rather than photons to study electronic structures of a material. The O_2 photoelectron spectrum emits two bands of 52.3898 MJ* mol^{-1} and 52.311 MJ*mol^{-1}. This emission corresponds to the 1\(\s) ionization of the oxygen electron. In your own words explain the observation.

9.28

Would the ionization energies of non-bonding 1s electrons be different for HCl and Cl_2?

9.28

First write out the ground-state electron configuration for the homonuclear diatomic molecule F_2. Then given the experimental ionization energies for a 1s fluorine electron for HF and F_2 as 66.981 and 67.217 $MJ \cdot mol^{-1}$, explain why even though the 1s electrons of fluorine are not involved in the chemical bonds their ionization energies are different.

9.30

Try to solve for the ground-state term symbols for the diatomic molecules H_2^+ and H_2, given that the ground-state electron configuration for H_2^+ is $(\sigma_g 1s)^1$ and for H_2 is $(\sigma_g 1s)^2$.

9.30

Find the ground-state term symbols for B_2.

9.31

Solve for the ground-state molecular term symbols for O_2, N_2, N_2^+, and O_2^+.

9.E: The Chemical Bond: Diatomic Molecules (Exercises) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

MathChapters

A: Complex Numbers
B: Probability and Statistics
C: Vectors
D: Spherical Coordinates
E: Determinants
F: Matrices
G: Numerical Methods
H: Partial Differentiation
I: Series and Limits
J: The Binomial Distribution and Stirling's Appromixation

MathChapters is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

A: Complex Numbers

Real Numbers

Let us think of the ordinary numbers as set out on a line which goes to infinity in both positive and negative directions. We could start by taking a stretch of the line near the origin (that is, the point representing the number zero) and putting in the integers as follows:

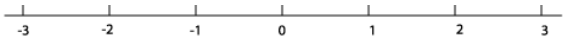


Figure $A.1$: Integers on number line

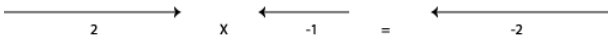
Next, we could add in rational numbers, such as ½, 23/11, etc., then the irrationals like $\sqrt{2}$, then numbers like π, and so on, so any number you can think of has its place on this line. Now let's take a slightly different point of view, and think of the numbers as represented by a *vector* from the origin to that number, so 1 is



and, for example, –2 is represented by:



Note that if a number is multiplied by –1, the corresponding vector is turned through 180 degrees. In pictures,



The "vector" 2 is turned through π, or 180 degrees, when you multiply it by –1.

Example $A.1$

What are the square roots of 4?

Solving Quadratic Equations

In solving the standard quadratic equation

$$ax^2 + bx + c = 0 \tag{A.1}$$

we find the solution to be:

$$x = \frac{-b \pm \sqrt{b^2 - ac}}{2a} \tag{A.2}$$

The problem with this is that sometimes the expression inside the square root is negative. What does that signify? For some problems in physics, it means there is no solution. For example, if I throw a ball directly upwards at 10 meters per sec, and ask when will it reach a height of 20 meters, taking g = 10 m per sec^2, the solution of the quadratic equation for the time t has a negative number inside the square root, and that means that the ball doesn't get to 20 meters, so the question didn't really make sense.

We shall find, however, that there are other problems, in wide areas of physics, where negative numbers inside square roots have an important physical significance. For that reason, we need to come up with a scheme for interpreting them.

The simplest quadratic equation that gives trouble is:

$$x^2 + 1 = 0 \tag{A.3}$$

the solutions being

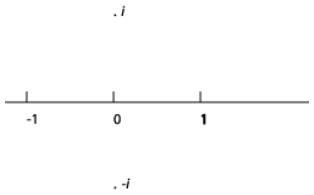
$$x = \pm\sqrt{-1} \tag{A.4}$$

What does that mean? We've just seen that the square of a positive number is positive, and the square of a negative number is also positive, since multiplying one negative number, which points backwards, by another, which turns any vector through 180°, gives a

positive vector. Another way of saying the same thing is to regard the minus sign itself, -, as an operator which turns the number it is applied to through 180º. Now $(-2) \times (-2)$ has two such rotations in it, giving the full 360º back to the positive axis.

To make sense of the square root of a negative number, we need to find something which when multiplied by itself gives a negative number. Let's concentrate for the moment on the square root of –1, from the quadratic equation above. Think of –1 as the operator – acting on the vector 1, so the – turns the vector through 180º. We need to find the square root of this operator, the operator which applied *twice* gives the rotation through 180º. Put like that, it is pretty obvious that the operator we want rotates the vector 1 through 90º.

But if we take a positive number, such as 1, and rotate its vector through 90º only, it isn't a number at all, at least in our original sense, since we put all known numbers on one line, and we've now rotated 1 away from that line. The new number created in this way is called a pure imaginary number, and is denoted by i.



Once we've found the square root of –1, we can use it to write the square root of any other negative number—for example, $2i$ is the square root of -4. Putting together a real number from the original line with an imaginary number (a multiple of i) gives a *complex number*. Evidently, complex numbers fill the entire two-dimensional plane. Taking ordinary Cartesian coordinates, any point P in the plane can be written as (x, y) where the point is reached from the origin by going x units in the direction of the positive real axis, then y units in the direction defined by i, in other words, the y axis.

Thus the point P with coordinates (x, y) can be identified with the complex number z, where

$$z = x + iy. \tag{A.5}$$

The plane is often called the *complex plane*, and representing complex numbers in this way is sometimes referred to as an **Argand Diagram** (Figure $A.2$).

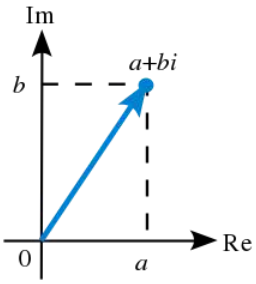


Figure $A.2$: A complex number can be visually represented as a pair of numbers (a, b) forming a vector on a diagram called an Argand diagram, representing the complex plane. "Re" is the real axis, "Im" is the imaginary axis, and i satisfies $i^2 = -1$. Image used with permission (Cc BY-SA 4.0; Wolfkeeper).

Visualizing the complex numbers as two-dimensional vectors, it is clear how to *add* two of them together. If $z_1 = x_1 + iy_1$, and $z_2 = x - 2 + iy_2$, then $z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$. The real parts and imaginary parts are added separately, just like vector components.

Multiplying two complex numbers together does not have quite such a simple interpretation. It is, however, quite straightforward—ordinary algebraic rules apply, with i^2 replaced where it appears by -1. So for example, to multiply $z_1 = x_1 + iy_1$ by $z_2 = x_2 + iy_2$,

$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) \tag{A.6}$$

$$= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1). \tag{A.7}$$

Polar Coordinates

Some properties of complex numbers are most easily understood if they are represented by using the polar coordinates r, θ instead of (x, y) to locate z in the complex plane (Figure $A.3$).

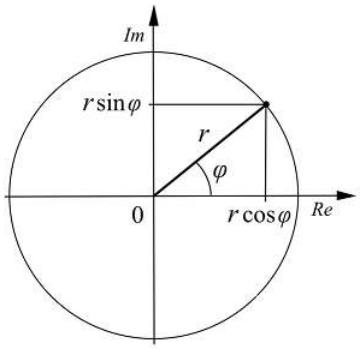


Figure $A.3$: Polar Argand diagram. Image used with permission (CC BY-SA 4.0; Svjo).

Note that $z = x + iy$ can be written $r(\cos\theta + i\sin\theta)$ from Figure $A.3$. In fact, this representation leads to a clearer picture of multiplication of two complex numbers:

$$z_1 z_2 = [r_2(\cos(\theta_1 + i\sin\theta_1)][r_2(\cos(\theta_2 + i\sin\theta_2)] \tag{A.8}$$

$$= r_1 r_2\,[(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2)] \tag{A.9}$$

$$= r_1 r_2\,[\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)] \tag{A.10}$$

So, if

$$z = r(\cos\theta + i\sin\theta) \tag{A.11}$$

$$= z_1 z_2 \tag{A.12}$$

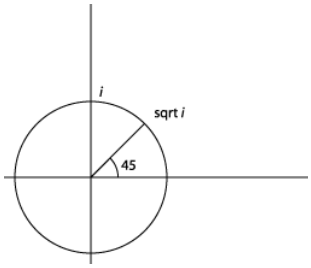
then

$$r = r_1 r_2 \tag{A.13}$$

and

$$\theta = \theta_1 \theta_2 \tag{A.14}$$

That is to say, to multiply together two complex numbers, we *multiply* the r's – called the *moduli* – and *add* the phases, the θ's. The modulus r is often denoted by \(|z|\), and called **mod** z, the phase θ is sometimes referred to as **arg** z. For example, $|i| = 1$, $\arg i = \pi/2$.

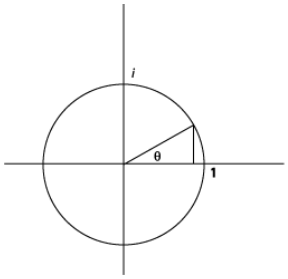


We can now see that, although we had to introduce these complex numbers to have a $\sqrt{-1}$, we do not need to bring in new types of numbers to get $\sqrt{-1}$, or $\sqrt{i}$. Clearly, $|\sqrt{i}| = 1$, $arg\sqrt{i} = 45°$. It is on the circle of unit radius centered at the origin, at 45°, and squaring it just doubles the angle.

The Unit Circle

In fact this circle—called the *unit circle*—plays an important part in the theory of complex numbers and every point on the circle has the form

$$z = \cos\theta + i\sin\theta = Cis(\theta) \tag{A.15}$$



Since all points on the unit circle have $|z| = 1$, by definition, multiplying any two of them together just amounts to adding the angles, so our new function $Cis(\theta)$ satisfies

$$Cis(\theta_1)Cis(\theta_2) = Cis(\theta_1 + \theta_2). \tag{A.16}$$

But that is just how multiplication works for exponents! That is,

$$a^{\theta_1}a^{\theta_2} = a^{\theta_1+\theta_2} \tag{A.17}$$

for a any constant, which strongly suggests that maybe our function $Cis(\theta$ is nothing but some constant a raised to the power θ, that is,

$$Cis(\theta) = a^{\theta} \tag{A.18}$$

It turns out to be convenient to write $a^{\theta} = e^{\ln a)\theta} = e^{A\theta}$, where $A = \ln a$. This line of reasoning leads us to write

$$\cos\theta + i\sin\theta = e^{A\theta} \tag{A.19}$$

Now, for the above "addition formula" to work for multiplication, A must be a constant, *independent* of θ. Therefore, we can find the value of A by choosing θ for which things are simple. We take θ to be very small—in this limit

$$\cos\theta \approx 1$$
$$\sin\theta \approx \theta$$

and

$$e^{A\theta} \approx 1 + A\theta \tag{A.20}$$

after dropping terms of order θ^2 and higher.

Substituting these values into Equation A.19 givesθ

So we find:

$$(\cos\theta + i\sin\theta)e^{i\theta} \tag{A.21}$$

To test this result, we expand $e^{i\theta}$:

$$e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \ldots \tag{A.22}$$
$$= 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{i\theta^5}{5!} + \ldots \tag{A.23}$$
$$= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!}\right) + i\left(\theta - \frac{i\theta^3}{3!} + \frac{i\theta^5}{5!}\right) \tag{A.24}$$
$$= \cos\theta + i\sin\theta \tag{A.25}$$

We write $= \cos\theta + i\sin\theta$ in the last line because the series in the brackets are precisely the Taylor series for $\cos\theta$ and $\sin\theta$ confirming our equation for $e^{i\theta}$. Changing the sign of θ it is easy to see that

$$e^{-i\theta} = \cos\theta - i\sin\theta \tag{A.26}$$

so the two trigonometric functions can be expressed in terms of exponentials of complex numbers:

$$\cos(\theta) = \frac{1}{2}\left(e^{i\theta} + e^{-i\theta}\right) \tag{A.27}$$

$$\sin(\theta) = \frac{1}{2i}\left(e^{i\theta} - e^{-i\theta}\right) \tag{A.28}$$

Euler Formula

The Euler formula states that any complex number can be written (Equation A.25):

$$e^{i\theta} = \cos\theta + i\sin\theta$$

Contributors

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

A: Complex Numbers is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

B: Probability and Statistics

A random variable X can have more than one value x as an outcome. Which value the variable has in a particular case is a matter of *chance* and cannot be predicted other than that we associate a probability to the outcome. **Probability** p is a number between 0 and 1 that indicates the likelihood that the variable X has a particular outcome x. The set of outcomes and their probabilities form a ***probability distribution***. There are two kinds of distributions:

1. discrete ones
2. continuous ones

Total probability should always add up to unity.

Discrete Distributions

A good example of a discrete distribution is a *true coin*. The random variable X can have two values:

1. heads (0)
2. tails (1)

Both have equal probability and as the sum must equal unity, the probability must be ½ for each. 'The probability that X=heads' is written formally as:

$$Pr(X = heads) = Pr(X = 0) = 0.5$$

The random function is written as a combination of three statements.

- Pr(X=0) = ½
- Pr(X=1) = ½
- elsewhere Pr = 0

Continuous Distributions

Now consider a spherical die. One could say it has an infinite number of facets that it can land on. Thus the number of outcomes $n = \infty$, this make each probability

$$p = 1/\infty = 0.$$

This creates a bit of a mathematical problem, because how can we get a total probability of unity by adding up zeros? Also, if we divide the sphere in a northern and a southern hemisphere clearly the probability that it lands on a point in say, the north should be ½. Still, p = 0 for all points. We introduce a new concept: **probability density** over which we *integrate* rather than *sum*. We assign an equal density to each point of the sphere and make sure that if we integrate over a hemisphere we get ½. (This involves two angles θ and φ and an integration over them and I won't go into that).

A bit simpler example of a continuous distribution than the spherical die is the 1D ***uniform*** distribution. It is the one that the Excel function =RAND() produces to good approximation. Its probability density is defined as

- f(x) = 1 for 0<1>
- f(x) = 0 elsewhere

The figure shows a (bivariate) uniform distribution.

The probability that the outcome is smaller than 0.5 is written as Pr(X<0.5) and is found by integrating from 0 to 0.5 over f(x).

Notice that in each individual outcome b the probability is indeed zero because an integral from b to b is always zero, even if the probability density f(b) is not zero. Clearly the probability and the probability density are not the same. Unfortunately the distinction between probability and probability density is often not properly made in the physical sciences. Moments can also be computed for continuous distributions by integrating over the probability density

Another well-known continuous distribution is the ***normal*** (or ***Gaussian***) distribution, defined as:

\[f(x) = 1/[√(2π)σ] * exp(-½[(x-μ)/σ]^2) \nonumber \]

(Notice the normalization factor 1/[√(2π)σ])

We can also compute moments of continuous distribution. Instead of using a summation we now have to evaluate an integral:

$$\langle X \rangle = \int [f(x)^* x] dx$$

$$\langle X^2 \rangle = \int [f(x)^* x^2] dx$$

For the normal distribution $\langle X \rangle = \mu$

? Exercise

Compute $^2>$ and $^3>$ for the uniform distribution. answer>

Indistinguishable Outcomes

When flipping two coins we could get four outcomes: two heads (0), heads plus tails (1), tails plus heads (1), two tails (2)

Each outcome is equally likely, this implies a probability of ¼ for each:

The probability of a particular outcome is often abbreviated simply to p. The two middle outcomes collapse into one with p=¼+¼= ½ if the coins are **indistinguishable**. We will see that this concept has very important consequences in statistical thermodynamics.

If we cannot distinguish the two outcomes leading to X_{tot}=1 we get the following random function:

Notice that it is quite possible to have a distribution where the probabilities differ from outcome to outcome. Often the p values are given as f(x), a function of x. An example:

X3 defined as:

The factor 1/6 makes sure the probabilities add up to unity. Such a factor is known as a **normalization factor**. Again this concept is of prime importance in statistical thermodynamics.

Another example of a discrete distribution is a die. If it has 6 sides (the most common die) there are six outcomes, each with p= 1/6. There are also dice with n=4, 12 or 20 sides. Each outcome will then have p= 1/n.

Moments of Distributions

In important aspect of probability distributions are the **moments** of the distribution. They are values computed by summing over the whole distribution.

The **zero** order moment is simply the sum of all p and that is unity:

The **first** moment multiplies each outcome with its probability and sums over all outcomes:

This moment is known as the ***average*** or ***mean***. (It is what we have done to your grades for years...)

For one coin is ½, for two coins is 1. (Exercise: verify this)

The **second** moment is computed by summing the product of the square of X and p:

The notation is used a lot in quantum mechanics, often in the form <ψ*ψ> or <ψ*|h|ψ>. The <.. part is known as the ***bra***, the ..> part as the ***ket***. (Together ***bra***(c)***ket***)

Intermezzo: The strange employer

You have a summer job but your employer likes games of chance. At the end of every day he rolls a die and pays you the square of the outcome in dollars per hour. So on a lucky day you'd make $36.- per hour, but on a bad day $1.-. Is this a bad deal? What would you make on the average over a longer period?

To answer this we must compute the second moment $^2>$ of the distribution:

(I have taken p=1/6 out of brackets because the value is the same for all six outcomes)

As you see in the intermezzo, the value of the second moment is in this case what you ***expect*** to be making on the long term. Moments are examples of what is know as **expectation values**. Another term you may run into that of a **functional**. A functional is a *number* computed by some operation (such as summation or integration) over a whole function. Moments are clearly an example of that too.

B: Probability and Statistics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **32.2: Probability and Statistics** is licensed CC BY-NC-SA 4.0.

C: Vectors

In this chapter we will review a few concepts you probably know from your physics courses. This chapter does not intend to cover the topic in a comprehensive manner, but instead touch on a few concepts that you will use in your physical chemistry classes.

A vector is a quantity that has both a magnitude and a direction, and as such they are used to specify the position, velocity and momentum of a particle, or to specify a force. Vectors are usually denoted by boldface symbols (e.g. $\mathbf{u}$) or with an arrow above the symbol (e.g. $\vec{u}$). A tilde placed above or below the name of the vector is also commonly used in shorthand ($\tilde{u}$,$\underset{\sim}{u}$).

If we multiply a number a by a vector $\mathbf{v}$, we obtain a new vector that is parallel to the original but with a length that is a times the length of $\mathbf{v}$. If a is negative $a\mathbf{v}$ points in the opposite direction than $\mathbf{v}$. We can express any vector in terms of the so-called unit vectors. These vectors, which are designated $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$, have unit length and point along the positive x, y and z axis of the cartesian coordinate system (Figure $C.1$). The symbol $\hat{\mathbf{i}}$ is read "i-hat". Hats are used to denote that a vector has unit length.

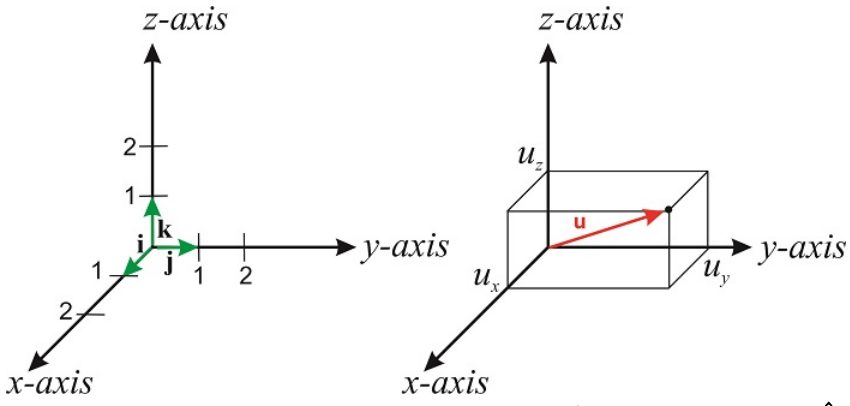


Figure $C.1$: Left: The unit vectors. Right: A vector $\mathbf{u}$ can be expressed in terms of the unit vectors as $\mathbf{u} = u_x\hat{\mathbf{i}} + u_y\hat{\mathbf{j}} + u_z\hat{\mathbf{k}}$ (CC BY-NC-SA; Marcia Levitus)

The length of $\mathbf{u}$ is its magnitude (or modulus), and is usually denoted by u:

$$u = |u| = (u_x^2 + u_y^2 + u_z^2)^{1/2} \tag{C.1}$$

If we have two vectors $\mathbf{u} = u_x\hat{\mathbf{i}} + u_y\hat{\mathbf{j}} + u_z\hat{\mathbf{k}}$ and $\mathbf{v} = v_x\hat{\mathbf{i}} + v_y\hat{\mathbf{j}} + v_z\hat{\mathbf{k}}$, we can add them to obtain

$$\mathbf{u} + \mathbf{v} = (u_x + v_x)\hat{\mathbf{i}} + (u_y + v_y)\hat{\mathbf{j}} + (u_z + v_z)\hat{\mathbf{k}}$$

or subtract them to obtain:

$$\mathbf{u} - \mathbf{v} = (u_x - v_x)\hat{\mathbf{i}} + (u_y - v_y)\hat{\mathbf{j}} + (u_z - v_z)\hat{\mathbf{k}}$$

When it comes to multiplication, we can perform the product of two vectors in two different ways. The first, which gives a scalar (a number) as the result, is called scalar product or dot product. The second, which gives a vector as a result, is called the vector (or cross) product. Both are important operations in physical chemistry.

The Scalar Product

The scalar product of vectors $\mathbf{u}$ and $\mathbf{v}$, also known as the dot product or inner product, is defined as (notice the dot between the symbols representing the vectors)

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}||\mathbf{v}| \cos\theta$$

where θ is the angle between the vectors. Notice that the dot product is zero if the two vectors are perpendicular to each other, and equals the product of their absolute values if they are parallel. It is easy to prove that

$$\mathbf{u} \cdot \mathbf{v} = u_x v_x + u_y v_y + u_z v_z$$

Example C.1

Show that the vectors

$$\mathbf{u_1} = \frac{1}{\sqrt{3}}\hat{\mathbf{i}} + \frac{1}{\sqrt{3}}\hat{\mathbf{j}} + \frac{1}{\sqrt{3}}\hat{\mathbf{k}}$$
$$\mathbf{u_2} = \frac{1}{\sqrt{6}}\hat{\mathbf{i}} - \frac{2}{\sqrt{6}}\hat{\mathbf{j}} + \frac{1}{\sqrt{6}}\hat{\mathbf{k}}$$
$$\mathbf{u_3} = -\frac{1}{\sqrt{2}}\hat{\mathbf{i}} + \frac{1}{\sqrt{2}}\hat{\mathbf{k}}$$

are of unit length and are mutually perpendicular.

Solution

The length of the vectors are:

$$|\mathbf{u_1}| = \left[\left(\frac{1}{\sqrt{3}}\right)^2 + \left(\frac{1}{\sqrt{3}}\right)^2 + \left(\frac{1}{\sqrt{3}}\right)^2\right]^{1/2} = \left[\frac{1}{3} + \frac{1}{3} + \frac{1}{3}\right]^{1/2} = 1$$
$$|\mathbf{u_2}| = \left[\left(\frac{1}{\sqrt{6}}\right)^2 + \left(-\frac{2}{\sqrt{6}}\right)^2 + \left(\frac{1}{\sqrt{6}}\right)^2\right]^{1/2} = \left[\frac{1}{6} + \frac{4}{6} + \frac{1}{6}\right]^{1/2} = 1$$
$$|\mathbf{u_3}| = \left[\left(-\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2\right]^{1/2} = \left[\frac{1}{2} + \frac{1}{2}\right]^{1/2} = 1$$

To test if two vectors are perpendicular, we perform the dot product:

$$\mathbf{u_1} \cdot \mathbf{u_2} = \left(\frac{1}{\sqrt{3}}\frac{1}{\sqrt{6}} - \frac{1}{\sqrt{3}}\frac{2}{\sqrt{6}} + \frac{1}{\sqrt{3}}\frac{1}{\sqrt{6}}\right) = 0$$
$$\mathbf{u_1} \cdot \mathbf{u_3} = \left(-\frac{1}{\sqrt{3}}\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}}\frac{1}{\sqrt{2}}\right) = 0$$
$$\mathbf{u_2} \cdot \mathbf{u_3} = \left(-\frac{1}{\sqrt{6}}\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{6}}\frac{1}{\sqrt{2}}\right) = 0$$

Therefore, we just proved that the three pairs are mutually perpendicular, and the three vectors have unit length. In other words, these vectors are the vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ rotated in space.

If the dot product of two vectors (of any dimension) is zero, we say that the two vectors are orthogonal. If the vectors have unit length, we say they are normalized. If two vectors are both normalized and they are orthogonal, we say they are orthonormal. The set of vectors shown in the previous example form an orthonormal set.[vectors:orthonormal] These concepts also apply to vectors that contain complex entries, but how do we perform the dot product in this case?

In general, the square of the modulus of a vector is

$$|\mathbf{u}|^2 = \mathbf{u} \cdot \mathbf{u} = u_x^2 + u_y^2 + u_z^2.$$

However, this does not work correctly for complex vectors. The square of i is -1, meaning that we risk having non-positive absolute values. To address this issue, we introduce a more general version of the dot product:

$$\mathbf{u} \cdot \mathbf{v} = u_x^* v_x + u_y^* v_y + u_z^* v_z,$$

where the "$*$" refers to the complex conjugate. Therefore, to calculate the modulus of a vector $\mathbf{u}$ that has complex entries, we use its complex conjugate:

$$|\mathbf{u}|^2 = \mathbf{u}^* \cdot \mathbf{u}$$

Example C.2: Calculating the Modulus of a vector

Calculate the modulus of the following vector:

$$\mathbf{u} = \hat{\mathbf{i}} + i\hat{\mathbf{j}}$$

Solution

$$|\mathbf{u}|^2 = \mathbf{u}^* \cdot \mathbf{u} = (\hat{\mathbf{i}} - i\hat{\mathbf{j}})(\hat{\mathbf{i}} + i\hat{\mathbf{j}}) = (1)(1) + (-i)(i) = 2 \rightarrow |\mathbf{u}| = \sqrt{2}$$

Analogously, if vectors contain complex entries, we can test whether they are orthogonal or not by checking the dot product $\mathbf{u}^* \cdot \mathbf{v}$.

Example C.3: Confirming orthogonality

Determine if the following pair of vectors are orthogonal (do not confuse the irrational number i with the unit vector $\hat{\mathbf{i}}$!)

$$\mathbf{u} = \hat{\mathbf{i}} + (1 - i)\hat{\mathbf{j}}$$

and

$$\mathbf{v} = (1 + i)\hat{\mathbf{i}} + \hat{\mathbf{j}}$$

Solution

$$\mathbf{u}^* \cdot \mathbf{v} = (\hat{\mathbf{i}} + (1 + i)\hat{\mathbf{j}})((1 + i)\hat{\mathbf{i}} + \hat{\mathbf{j}}) = (1)(1 + i) + (1 + i)(1) = 2 + 2i \neq 0$$

Therefore, the vectors are not orthogonal.

The Vector Product

The vector product of two vectors is a vector defined as

$$\mathbf{u} \times \mathbf{v} = |\mathbf{u}||\mathbf{v}|\mathbf{n} \sin\theta$$

where θ is again the angle between the two vectors, and $\mathbf{n}$ is the unit vector perpendicular to the plane formed by $\mathbf{u}$ and $\mathbf{v}$. The direction of the vector $\mathbf{n}$ is given by the right-hand rule. Extend your right hand and point your index finger in the direction of $\mathbf{u}$ (the vector on the left side of the $\times$ symbol) and your forefinger in the direction of $\mathbf{v}$. The direction of $\mathbf{n}$, which determines the direction of $\mathbf{u} \times \mathbf{v}$, is the direction of your thumb. If you want to revert the multiplication, and perform $\mathbf{v} \times \mathbf{u}$, you need to point your index finger in the direction of $\mathbf{v}$ and your forefinger in the direction of $\mathbf{u}$ (still using the right hand!). The resulting vector will point in the opposite direction (Figure C.1).

The magnitude of $\mathbf{u} \times \mathbf{v}$ is the product of the magnitudes of the individual vectors times $\sin\theta$. This magnitude has an interesting geometrical interpretation: it is the area of the parallelogram formed by the two vectors (Figure C.1).

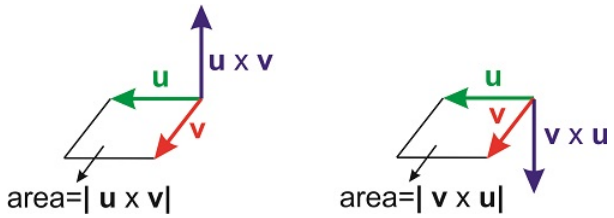


Figure C.1: The vector product (CC BY-NC-SA; Marcia Levitus)

The cross product can also be expressed as a determinant:

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix}$$

Example C.1:

Given $\mathbf{u} = -2\hat{\mathbf{i}} + \hat{\mathbf{j}} + \hat{\mathbf{k}}$ and $\mathbf{v} = 3\hat{\mathbf{i}} - \hat{\mathbf{j}} + \hat{\mathbf{k}}$, calculate $\mathbf{w} = \mathbf{u} \times \mathbf{v}$ and verify that the result is perpendicular to both $\mathbf{u}$ and $\mathbf{v}$.

Solution

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ -2 & 1 & 1 \\ 3 & -1 & 1 \end{vmatrix}$$

$$= \hat{\mathbf{i}}(1+1) - \hat{\mathbf{j}}(-2-3) + \hat{\mathbf{k}}(2-3)$$

$$= 2\hat{\mathbf{i}} + 5\hat{\mathbf{j}} - \hat{\mathbf{k}}$$

To verify that two vectors are perpendicular we perform the dot product:

$$\mathbf{u} \cdot \mathbf{w} = (-2)(2) + (1)(5) + (1)(-1) = 0$$

$$\mathbf{v} \cdot \mathbf{w} = (3)(2) + (-1)(5) + (1)(-1) = 0$$

An important application of the cross product involves the definition of the angular momentum. If a particle with mass m moves a velocity $\mathbf{v}$ (a vector), its (linear) momentum is $\mathbf{p} = m\mathbf{v}$. Let $\mathbf{r}$ be the position of the particle (another vector), then the angular momentum of the particle is defined as

$$\mathbf{l} = \mathbf{r} \times \mathbf{p}$$

The angular momentum is therefore a vector perpendicular to both $\mathbf{r}$ and $\mathbf{p}$. Because the position of the particle needs to be defined with respect to a particular origin, this origin needs to be specified when defining the angular momentum.

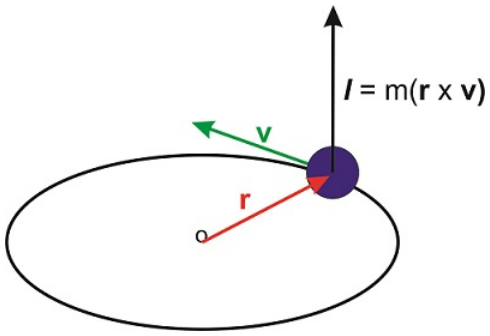


Figure C.2: The angular momentum of a particle of position $\mathbf{r}$ from the origin and momentum $\mathbf{p} = m\mathbf{v}$ (CC BY-NC-SA; Marcia Levitus)

Vector Normalization

A vector of any given length can be divided by its modulus to create a unit vector (i.e. a vector of unit length). We will see applications of unit (or normalized) vectors in the next chapter.

For example, the vector

$$\mathbf{u} = \hat{\mathbf{i}} + \hat{\mathbf{j}} + i\hat{\mathbf{k}}$$

has a magnitude:

$$|\mathbf{u}|^2 = 1^2 + 1^2 + (-i)(i) = 3 \rightarrow |\mathbf{u}| = \sqrt{3}$$

Therefore, to normalize this vector we divide all the components by its length:

$$\hat{\mathbf{u}} = \frac{1}{\sqrt{3}}\hat{\mathbf{i}} + \frac{1}{\sqrt{3}}\hat{\mathbf{j}} + \frac{i}{\sqrt{3}}\hat{\mathbf{k}}$$

Notice that we use the “hat” to indicate that the vector has unit length.

Need help? The links below contain solved examples.

Operations with vectors: http://tinyurl.com/mw4qmz8

External links:

- The dot product: http://patrickjmt.com/vectors-the-dot-product/
- The cross product: http://patrickjmt.com/the-cross-product/
- The dot and cross product: http://www.youtube.com/watch?v=enr7JqvehJs

---

C: Vectors is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **14.1: Introduction to Vectors** by Marcia Levitus is licensed CC BY-NC-SA 4.0. Original source: https://www.public.asu.edu/~mlevitus/chm240/book.pdf.
- **32.5: Vectors** by Marcia Levitus is licensed CC BY-NC-SA 4.0. Original source: https://www.public.asu.edu/~mlevitus/chm240/book.pdf.
- **14.2: The Scalar Product** by Marcia Levitus is licensed CC BY-NC-SA 4.0. Original source: https://www.public.asu.edu/~mlevitus/chm240/book.pdf.
- **14.3: The Vector Product** by Marcia Levitus is licensed CC BY-NC-SA 4.0. Original source: https://www.public.asu.edu/~mlevitus/chm240/book.pdf.
- **14.4: Vector Normalization** by Marcia Levitus is licensed CC BY-NC-SA 4.0. Original source: https://www.public.asu.edu/~mlevitus/chm240/book.pdf.

D: Spherical Coordinates

Objectives

- Understand the concept of area and volume elements in cartesian, polar and spherical coordinates.
- Be able to integrate functions expressed in polar or spherical coordinates.
- Understand how to normalize orbitals expressed in spherical coordinates, and perform calculations involving triple integrals.

Coordinate Systems

The simplest coordinate system consists of coordinate axes oriented perpendicularly to each other. These coordinates are known as cartesian coordinates or rectangular coordinates, and you are already familiar with their two-dimensional and three-dimensional representation. In the plane, any point P can be represented by two signed numbers, usually written as (x, y), where the coordinate x is the distance perpendicular to the x axis, and the coordinate y is the distance perpendicular to the y axis (Figure $D.1$, left). In space, a point is represented by three signed numbers, usually written as (x, y, z) (Figure $D.1$, right).

Often, positions are represented by a vector, $\vec{r}$, shown in red in Figure $D.1$. In three dimensions, this vector can be expressed in terms of the coordinate values as $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$, where $\hat{i} = (1, 0, 0)$, $\hat{j} = (0, 1, 0)$ and $\hat{z} = (0, 0, 1)$ are the so-called unit vectors.

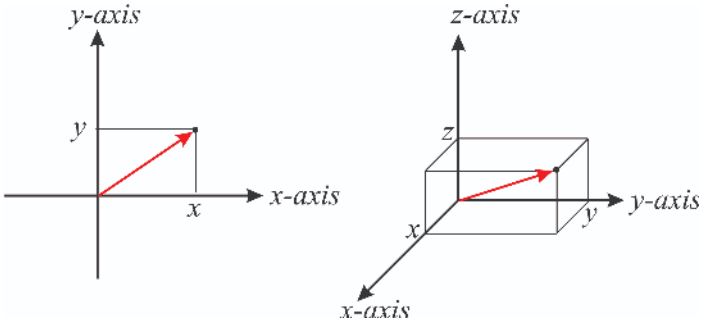


Figure $D.1$: Cartesian coordinates (CC BY-NC-SA; Marcia Levitus)

We already know that often the symmetry of a problem makes it natural (and easier!) to use other coordinate systems. In two dimensions, the polar coordinate system defines a point in the plane by two numbers: the distance r to the origin, and the angle θ that the position vector forms with the x-axis. Notice the difference between $\vec{r}$, a vector, and r, the distance to the origin (and therefore the modulus of the vector). Vectors are often denoted in bold face (e.g. $\mathbf{r}$) without the arrow on top, so be careful not to confuse it with r, which is a scalar.

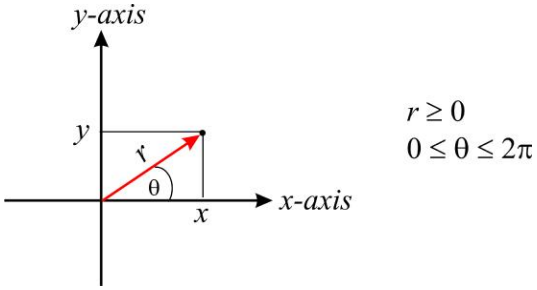


Figure $D.2$: Plane polar coordinates (CC BY-NC-SA; Marcia Levitus)

While in cartesian coordinates x, y (and z in three-dimensions) can take values from $-\infty$ to ∞, in polar coordinates r is a positive value (consistent with a distance), and θ can take values in the range $[0, 2\pi]$.

The relationship between the cartesian and polar coordinates in two dimensions can be summarized as:

$$x = r\cos\theta \tag{D.1}$$

$$y = r\sin\theta \tag{D.2}$$

$$r^2 = x^2 + y^2 \tag{D.3}$$

$$\tan\theta = y/x \tag{D.4}$$

In three dimensions, the spherical coordinate system defines a point in space by three numbers: the distance r to the origin, a polar angle ϕ that measures the angle between the positive x-axis and the line from the origin to the point P projected onto the xy-plane, and the angle θ defined as the is the angle between the z-axis and the line from the origin to the point P:

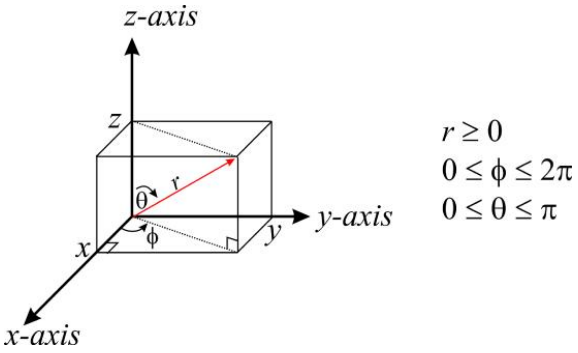


Figure $D.3$: Spherical coordinates (CC BY-NC-SA; Marcia Levitus)

Before we move on, it is important to mention that depending on the field, you may see the Greek letter θ (instead of ϕ) used for the angle between the positive x-axis and the line from the origin to the point P projected onto the xy-plane. That is, θ and ϕ may appear interchanged. This can be very confusing, so you will have to be careful. When using spherical coordinates, it is important that you see how these two angles are defined so you can identify which is which.

Spherical coordinates are useful in analyzing systems that are symmetrical about a point. For example a sphere that has the cartesian equation $x^2 + y^2 + z^2 = R^2$ has the very simple equation $r = R$ in spherical coordinates. Spherical coordinates are the natural coordinates for physical situations where there is spherical symmetry (e.g. atoms). The relationship between the cartesian coordinates and the spherical coordinates can be summarized as:

$$x = r\sin\theta\cos\phi \tag{D.5}$$

$$y = r\sin\theta\sin\phi \tag{D.6}$$

$$z = r\cos\theta \tag{D.7}$$

These relationships are not hard to derive if one considers the triangles shown in Figure $D.4$:

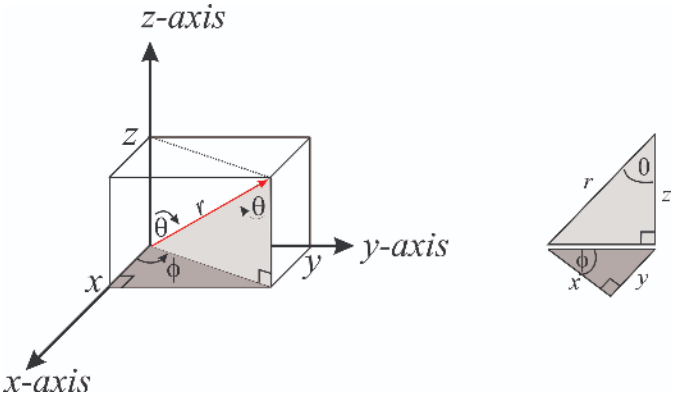


Figure $D.4$: Spherical coordinates (CC BY-NC-SA; Marcia Levitus)

Area and Volume Elements

In any coordinate system it is useful to define a differential area and a differential volume element. In cartesian coordinates the differential area element is simply $dA = dx\,dy$ (Figure $D.1$), and the volume element is simply $dV = dx\,dy\,dz$.

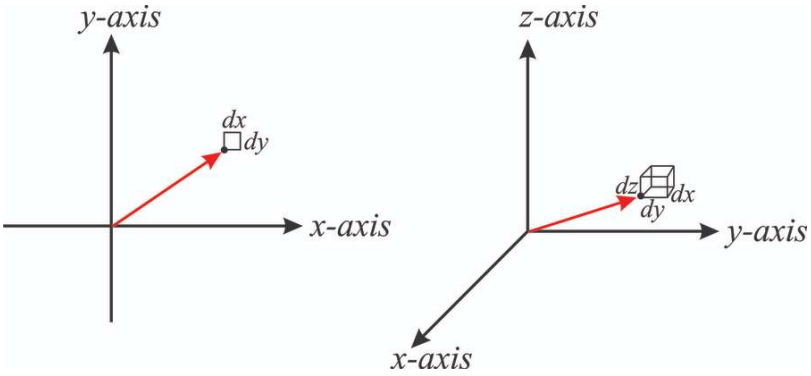


Figure D.1: Area and volume elements in cartesian coordinates (CC BY-NC-SA; Marcia Levitus)

We already performed double and triple integrals in cartesian coordinates, and used the area and volume elements without paying any special attention. For example, in example [c2v:c2vex1], we were required to integrate the function $|\psi(x,y,z)|^2$ over all space, and without thinking too much we used the volume element $dx\ dy\ dz$ (see page). We also knew that "all space" meant $-\infty \leq x \leq \infty$, $-\infty \leq y \leq \infty$ and $-\infty \leq z \leq \infty$, and therefore we wrote:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\psi(x,y,z)|^2\ dx\ dy\ dz = 1$$

But what if we had to integrate a function that is expressed in spherical coordinates? Would we just replace $dx\ dy\ dz$ by $dr\ d\theta\ d\phi$? The answer is no, because the volume element in spherical coordinates depends also on the actual position of the point. This will make more sense in a minute. Coming back to coordinates in two dimensions, it is intuitive to understand why the area element in cartesian coordinates is $dA = dx\ dy$ independently of the values of x and y. This is shown in the left side of Figure D.2. However, in polar coordinates, we see that the areas of the gray sections, which are both constructed by increasing r by dr, and by increasing θ by $d\theta$, depend on the actual value of r. Notice that the area highlighted in gray increases as we move away from the origin.

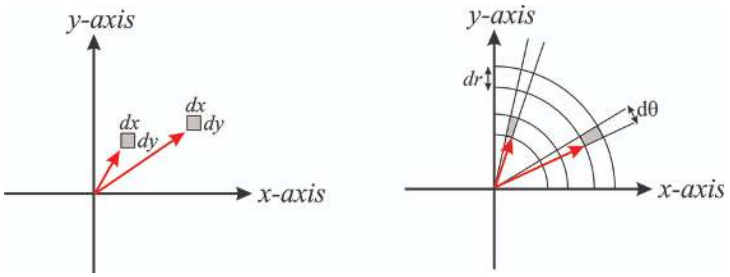


Figure D.2: Differential of area in cartesian and polar coordinates (CC BY-NC-SA; Marcia Levitus)

The area shown in gray can be calculated from geometrical arguments as

$$dA = \left[\pi(r+dr)^2 - \pi r^2\right]\frac{d\theta}{2\pi}.$$

Because $dr << 0$, we can neglect the term $(dr)^2$, and $dA = r\ dr\ d\theta$ (see Figure 10.2.3).

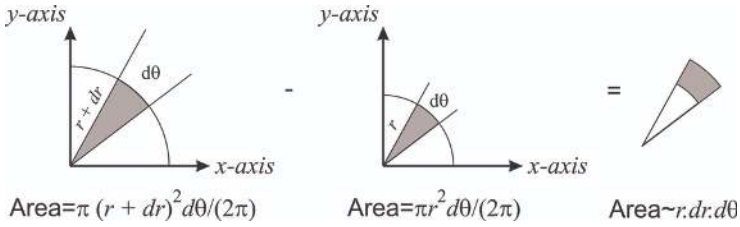


Figure D.3: Differential of area in polar coordinates (CC BY-NC-SA; Marcia Levitus)

Let's see how this affects a double integral with an example from quantum mechanics. The wave function of the ground state of a two dimensional harmonic oscillator is: $\psi(x,y) = Ae^{-a(x^2+y^2)}$. We know that the quantity $|\psi|^2$ represents a probability density, and as such, needs to be normalized:

$$\int_{all\ space} |\psi|^2\, dA = 1$$

This statement is true regardless of whether the function is expressed in polar or cartesian coordinates. However, the limits of integration, and the expression used for dA, will depend on the coordinate system used in the integration.

In cartesian coordinates, "all space" means $-\infty < x < \infty$ and $-\infty < y < \infty$. The differential of area is $dA = dxdy$:

$$\int_{all\ space} |\psi|^2\, dA = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A^2 e^{-2a(x^2+y^2)}\, dxdy = 1$$

In polar coordinates, "all space" means $0 < r < \infty$ and $0 < \theta < 2\pi$. The differential of area is $dA = r\ drd\theta$. The function $\psi(x,y) = Ae^{-a(x^2+y^2)}$ can be expressed in polar coordinates as: $\psi(r,\theta) = Ae^{-ar^2}$

$$\int_{all\ space} |\psi|^2\, dA = \int_{0}^{\infty}\int_{0}^{2\pi} A^2 e^{-2ar^2} r\, d\theta dr = 1$$

Both versions of the double integral are equivalent, and both can be solved to find the value of the normalization constant (A) that makes the double integral equal to 1. In polar coordinates:

$$\int_0^{\infty}\int_0^{2\pi} A^2 e^{-2ar^2} r\, d\theta dr = A^2 \int_0^{\infty} e^{-2ar^2} r\, dr \int_0^{2\pi} d\theta = A^2 \times \frac{1}{4a} \times 2\pi = 1$$

Therefore[1], $A = \sqrt{2a/\pi}$. The same value is of course obtained by integrating in cartesian coordinates.

It is now time to turn our attention to triple integrals in spherical coordinates. In cartesian coordinates, the differential volume element is simply $dV = dx\, dy\, dz$, regardless of the values of x, y and z. Using the same arguments we used for polar coordinates in the plane, we will see that the differential of volume in spherical coordinates is not $dV = dr\, d\theta\, d\phi$. The geometrical derivation of the volume is a little bit more complicated, but from Figure $D.4$ you should be able to see that dV depends on r and θ, but not on ϕ. The volume of the shaded region is

$$dV = r^2 \sin\theta\, d\theta\, d\phi\, dr \tag{D.8}$$

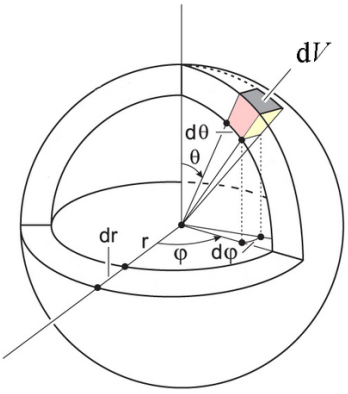


Figure $D.4$: Differential of volume in spherical coordinates (CC BY-NC-SA; Marcia Levitus)

We will exemplify the use of triple integrals in spherical coordinates with some problems from quantum mechanics. We already introduced the Schrödinger equation, and even solved it for a simple system in Section 5.4. We also mentioned that spherical coordinates are the obvious choice when writing this and other equations for systems such as atoms, which are symmetric around a point.

As we saw in the case of the particle in the box (Section 5.4), the solution of the Schrödinger equation has an arbitrary multiplicative constant. Because of the probabilistic interpretation of wave functions, we determine this constant by normalization. The same situation arises in three dimensions when we solve the Schrödinger equation to obtain the expressions that describe the possible states of the electron in the hydrogen atom (i.e. the orbitals of the atom). The Schrödinger equation is a partial differential

equation in three dimensions, and the solutions will be wave functions that are functions of r, θ and ϕ. The lowest energy state, which in chemistry we call the 1s orbital, turns out to be:

$$\psi_{1s} = Ae^{-r/a_0}$$

This particular orbital depends on r only, which should not surprise a chemist given that the electron density in all s-orbitals is spherically symmetric. We will see that p and d orbitals depend on the angles as well. Regardless of the orbital, and the coordinate system, the normalization condition states that:

$$\int_{all\ space} |\psi|^2\, dV = 1$$

For a wave function expressed in cartesian coordinates,

$$\int_{all\ space} |\psi|^2\, dV = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi^*(x,y,z)\psi(x,y,z)\, dxdydz$$

where we used the fact that $|\psi|^2 = \psi^*\psi$.

In spherical coordinates, "all space" means $0 \le r \le \infty$, $0 \le \phi \le 2\pi$ and $0 \le \theta \le \pi$. The differential dV is $dV = r^2 \sin\theta\, d\theta\, d\phi\, dr$, so

$$\int_{all\ space} |\psi|^2\, dV = \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \psi^*(r,\theta,\phi)\psi(r,\theta,\phi)\, r^2 \sin\theta\, drd\theta d\phi = 1$$

Let's see how we can normalize orbitals using triple integrals in spherical coordinates.

Example $D.1$

When solving the Schrödinger equation for the hydrogen atom, we obtain $\psi_{1s} = Ae^{-r/a_0}$, where A is an arbitrary constant that needs to be determined by normalization. Find A.

Solution

In spherical coordinates,

$$\int_{all\ space} |\psi|^2\, dV = \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \psi^*(r,\theta,\phi)\psi(r,\theta,\phi)\, r^2 \sin\theta\, drd\theta d\phi = 1$$

because this orbital is a real function, $\psi^*(r,\theta,\phi)\psi(r,\theta,\phi) = \psi^2(r,\theta,\phi)$. In this case, $\psi^2(r,\theta,\phi) = A^2e^{-2r/a_0}$.

Therefore,

$$\int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \psi^*(r,\theta,\phi)\psi(r,\theta,\phi)\, r^2 \sin\theta\, drd\theta d\phi = \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} A^2e^{-2r/a_0}\, r^2 \sin\theta\, drd\theta d\phi = 1$$

$$\int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} A^2e^{-2r/a_0}\, r^2 \sin\theta\, drd\theta d\phi = A^2 \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta \int_0^{\infty} e^{-2r/a_0}\, r^2\, dr$$

The result is a product of three integrals in one variable:

$$\int_0^{2\pi} d\phi = 2\pi$$

$$\int_0^{\pi} \sin\theta\, d\theta = -\cos\theta\Big|_0^{\pi} = 2$$

https://chem.libretexts.org/@go/page/198734

$$\int_0^\infty e^{-2r/a_0}\, r^2\, dr = ?$$

From the formula sheet:

$$\int_0^\infty x^n e^{-ax} dx = \frac{n!}{a^{n+1}},$$

where $a > 0$ and n is a positive integer.

In this case, $n = 2$ and $a = 2/a_0$, so:

$$\int_0^\infty e^{-2r/a_0}\, r^2\, dr = \frac{2!}{(2/a_0)^3} = \frac{2}{8/a_0^3} = \frac{a_0^3}{4}$$

Putting the three pieces together:

$$A^2 \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta \int_0^\infty e^{-2r/a_0}\, r^2\, dr = A^2 \times 2\pi \times 2 \times \frac{a_0^3}{4} = 1$$

$$A^2 \times \pi \times a_0^3 = 1 \rightarrow A = \frac{1}{\sqrt{\pi a_0^3}}$$

The normalized 1s orbital is, therefore:

$$\frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0}$$

This page titled D: Spherical Coordinates is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Marcia Levitus via source content that was edited to the style and standards of the LibreTexts platform.

- **32.6: Spherical Coordinates** is licensed CC BY-NC-SA 4.0.
- **10.1: Coordinate Systems** by Marcia Levitus is licensed CC BY-NC-SA 4.0. Original source: https://www.public.asu.edu/~mlevitus/chm240/book.pdf.
- **10.2: Area and Volume Elements** by Marcia Levitus is licensed CC BY-NC-SA 4.0. Original source: https://www.public.asu.edu/~mlevitus/chm240/book.pdf.

E: Determinants

The concept of determinants has its origin in the solution of simultaneous linear equations. In physical chemistry, they are an important tool in quantum mechanics

Suppose you want to solve the following system of two equations with two unknowns (x and y):

$$a_1x + b_1y = c_1$$

$$a_2x + b_2y = c_2$$

In order to find y, we could use the following general procedure: we multiply the first equation by a_2 and the second by a_1, and subtract one line from the other to cancel the term in x:

$$a_1x + b_1y = c_1 \overset{\times a_2}{\rightarrow} a_1a_2x + b_1a_2y = c_1a_2$$

$$a_2x + b_2y = c_2 \overset{\times a_1}{\rightarrow} a_1a_2x + b_2a_1y = c_2a_1$$

$$\left.\begin{aligned} a_1a_2x + b_1a_2y = c_1a_2 \\ a_1a_2x + b_2a_1y = c_2a_1 \end{aligned}\right\} \rightarrow (b_2a_1 - b_1a_2)y = a_1c_2 - a_2c_1 \rightarrow y = \frac{a_1c_2 - a_2c_1}{b_2a_1 - b_1a_2}$$

We can follow the same strategy to find x: we multiply the first equation by b_2 and the second by b_1, and subtract one line from the other to cancel the term in y:

$$a_1x + b_1y = c_1 \overset{\times b_2}{\rightarrow} a_1b_2x + b_1b_2y = c_1b_2$$

$$a_2x + b_2y = c_2 \overset{\times b_1}{\rightarrow} b_1a_2x + b_2b_1y = c_2b_1$$

$$\left.\begin{aligned} a_1b_2x + b_1b_2y = c_1b_2 \\ b_1a_2x + b_2b_1y = c_2b_1 \end{aligned}\right\} \rightarrow (b_2a_1 - b_1a_2)x = b_2c_1 - b_1c_2 \rightarrow x = \frac{b_2c_1 - b_1c_2}{b_2a_1 - b_1a_2}$$

We define a 2×2 determinant as:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb$$

The determinant, which is denoted with two parallel bars, is a number. For example,

$$\begin{vmatrix} 3 & -1 \\ 1/2 & 2 \end{vmatrix} = 3 \times 2 - (-1) \times 1/2 = 13/2$$

Let's look at the expressions we obtained for x and y, and write them in terms of determinants:

$$x = \frac{b_2c_1 - b_1c_2}{b_2a_1 - b_1a_2} = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$$

$$y = \frac{a_1c_2 - a_2c_1}{b_2a_1 - b_1a_2} = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$$

Let's look at our equations, and see how these determinants are constructed from the coefficients.

$$a_1x + b_1y = c_1$$

$$a_2x + b_2y = c_2$$

The determinant in the denominator of both x and y is the determinant of the coefficients on the left-side of the equal sign:

$$\left.\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}\right\} \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

The numerator in the expression of y is built by replacing the coefficients in the y-column with the coefficients on the right side of the equation:

$$\left.\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}\right\} \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$$

The numerator in the expression of x is built by replacing the coefficients in the x-column with the coefficients on the right side of the equation:

$$\left.\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}\right\} \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}$$

We can extend this idea to n equations with n unknowns $(x_1, x_2, x_3, ..., x_n)$.

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = b_1 \\ a_{21}x_1 & + & a_{22}c_2 & + & \cdots & + & a_{2n}x_n & = b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & \vdots \\ a_{n1}x_1 & + & a_{n2}c_2 & + & \cdots & + & a_{nn}x_n & = b_n \end{array}$$

Note that we use two subscripts to identify the coefficients. The first refers to the row, and the second to the column. Let's define the determinant D as the determinant of the coefficients of the equation (the ones on the left side of the equal sign):

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

and let's define the determinant D_k as the one obtained from D by replacement of the kth column of D by the column with elements $b_1, b_2 ... b_n$. For example, D_2 is

$$D_2 = \begin{vmatrix} a_{11} & b_1 & \cdots & a_{1n} \\ a_{21} & b_2 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & b_n & \cdots & a_{nn} \end{vmatrix}$$

The unknowns of the system of equations are calculated as:

$$x_1 = \frac{D_1}{D}, x_2 = \frac{D_2}{D}, ..., x_n = \frac{D_n}{D}$$

For example, let's say we want to find x, y and z in the following system of equations:

$$2x + 3y + 8z = 0$$

$$x - \frac{1}{2}y - 3z = \frac{1}{2}$$

$$-x - y - z = \frac{1}{2}$$

We can calculate the unknowns as;

$$x = \frac{D_1}{D}, y = \frac{D_2}{D}, z = \frac{D_3}{D}$$

where

$$D = \begin{vmatrix} 2 & 3 & 8 \\ 1 & -1/2 & -3 \\ -1 & -1 & -1 \end{vmatrix}$$

$$D_1 = \begin{vmatrix} 0 & 3 & 8 \\ 1/2 & -1/2 & -3 \\ 1/2 & -1 & -1 \end{vmatrix}$$

$$D_2 = \begin{vmatrix} 2 & 0 & 8 \\ 1 & 1/2 & -3 \\ -1 & 1/2 & -1 \end{vmatrix}$$

$$D_3 = \begin{vmatrix} 2 & 3 & 0 \\ 1 & -1/2 & 1/2 \\ -1 & -1 & 1/2 \end{vmatrix}$$

In order to do this, we need to learn how to solve 3×3 determinants, or in general, $n \times n$ determinants.

This page titled E: Determinants is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Marcia Levitus via source content that was edited to the style and standards of the LibreTexts platform.

- **13.1: The Solutions of Simultaneous Linear Equations** by Marcia Levitus is licensed CC BY-NC-SA 4.0. Original source: https://www.public.asu.edu/~mlevitus/chm240/book.pdf.

F: Matrices

Definitions

An $n \times m$ matrix is a two dimensional array of numbers with n rows and m columns. The integers n and m are called the *dimensions* of the matrix. If $n = m$ then the matrix is *square*. The numbers in the matrix are known as *matrix elements* (or just elements) and are usually given subscripts to signify their position in the matrix e.g. an element a_{ij} would occupy the ith row and j^{th} column of the matrix. For example:

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \tag{F.1}$$

is a 3×3 matrix with $a_{11} = 1$, $a_{12} = 2$, $a_{13} = 3$, $a_{21} = 4$ etc

In a square matrix, *diagonal* elements are those for which i=j (the numbers 1, 5, and 9 in the above example). *Off-diagonal* elements are those for which $i \neq j$ (2, 3, 4, 6, 7, and 8 in the above example). If all the off-diagonal elements are equal to zero then we have a *diagonal* matrix. We will see later that diagonal matrices are of considerable importance in group theory.

A *unit matrix* or *identity matrix* (usually given the symbol I) is a diagonal matrix in which all the diagonal elements are equal to 1. A unit matrix acting on another matrix has no effect – it is the same as the identity operation in group theory and is analogous to multiplying a number by 1 in everyday arithmetic.

The *transpose* A^T of a matrix A is the matrix that results from interchanging all the rows and columns. A *symmetric* matrix is the same as its transpose ($A^T = A$ i.e. $a_{ij} = a_{ji}$ for all values of i and j). The transpose of matrix M above (which is not symmetric) is

$$M^T = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \tag{F.2}$$

The sum of the diagonal elements in a square matrix is called the *trace* (or *character*) of the matrix (for the above matrix, the trace is $\chi = 1 + 5 + 9 = 15$). The traces of matrices representing symmetry operations will turn out to be of great importance in group theory.

A *vector* is just a special case of a matrix in which one of the dimensions is equal to 1. An $n \times 1$ matrix is a *column vector*; a $1 \times m$ matrix is a *row vector*. The components of a vector are usually only labeled with one index. A *unit vector* has one element equal to 1 and the others equal to zero (it is the same as one row or column of an identity matrix). We can extend the idea further to say that a single number is a matrix (or vector) of dimension 1×1.

Matrix Algebra

i. Two matrices with the same dimensions may be added or subtracted by adding or subtracting the elements occupying the same position in each matrix. e.g.

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 2 & 1 \\ 3 & 2 & 0 \end{pmatrix} \tag{F.3}$$

$$B = \begin{pmatrix} 2 & 0 & -2 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \tag{F.4}$$

$$A + B = \begin{pmatrix} 3 & 0 & 0 \\ 3 & 2 & 2 \\ 4 & 1 & 0 \end{pmatrix} \tag{F.5}$$

$$A - B = \begin{pmatrix} -1 & 0 & 4 \\ 1 & 2 & 0 \\ 2 & 3 & 0 \end{pmatrix} \tag{F.6}$$

ii. A matrix may be multiplied by a constant by multiplying each element by the constant.

$$4B = \begin{pmatrix} 8 & 0 & -8 \\ 4 & 0 & 4 \\ 4 & -4 & 0 \end{pmatrix} \tag{F.7}$$

$$3A = \begin{pmatrix} 3 & 0 & 6 \\ 6 & 6 & 3 \\ 9 & 6 & 0 \end{pmatrix} \tag{F.8}$$

iii. Two matrices may be multiplied together provided that the number of columns of the first matrix is the same as the number of rows of the second matrix i.e. an $n \times m$ matrix may be multiplied by an $m \times l$ matrix. The resulting matrix will have dimensions $n \times l$. To find the element a_{ij} in the product matrix, we take the dot product of row i of the first matrix and column j of the second matrix (i.e. we multiply consecutive elements together from row i of the first matrix and column j of the second matrix and add them together i.e. $c_{ij} = \Sigma_k\, a_{ik}b_{jk}$ e.g. in the 3×3 matrices A and B used in the above examples, the first element in the product matrix $C = AB$ is $c_{11} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}$

$$AB = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 2 & 1 \\ 3 & 2 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & -2 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & -2 & -2 \\ 7 & -1 & -2 \\ 8 & 0 & -4 \end{pmatrix} \tag{F.9}$$

An example of a matrix multiplying a vector is

$$A\mathbf{v} = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 2 & 1 \\ 3 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 9 \\ 7 \end{pmatrix} \tag{F.10}$$

> *Matrix multiplication is not generally commutative, a property that mirrors the behavior found earlier for symmetry operations within a point group.*

Direct Products

The direct product of two matrices (given the symbol $\otimes$) is a special type of matrix product that generates a matrix of higher dimensionality if both matrices have dimension greater than one. The easiest way to demonstrate how to construct a direct product of two matrices A and B is by an example:

$$A \otimes B = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \tag{F.11}$$

$$= \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} \tag{F.12}$$

$$= \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{21} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix} \tag{F.13}$$

Though this may seem like a somewhat strange operation to carry out, direct products crop up a great deal in group theory.

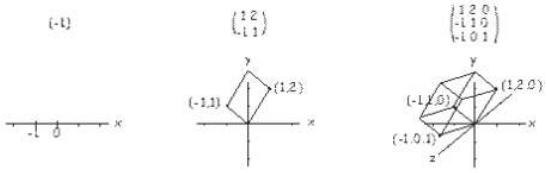
Inverse Matrices and Determinants

If two square matrices A and B multiply together to give the identity matrix I (i.e. $AB = I$) then B is said to be the *inverse* of A (written A^{-1}). If B is the inverse of A then A is also the inverse of B. Recall that one of the conditions imposed upon the symmetry operations in a group is that each operation must have an inverse. It follows by analogy that any matrices we use to represent symmetry elements must also have inverses. It turns out that a square matrix only has an inverse if its determinant is non-zero. For this reason (and others which will become apparent later on when we need to solve equations involving matrices) we need to learn a little about matrix determinants and their properties.

For every square matrix, there is a unique function of all the elements that yields a single number called the determinant. Initially it probably won't be particularly obvious why this number should be useful, but matrix determinants are of great importance both in

pure mathematics and in a number of areas of science. Historically, determinants were actually around before matrices. They arose originally as a property of a system of linear equations that 'determined' whether the system had a unique solution. As we shall see later, when such a system of equations is recast as a matrix equation this property carries over into the matrix determinant.

There are two different definitions of a determinant, one geometric and one algebraic. In the geometric interpretation, we consider the numbers across each row of an $n \times n$ matrix as coordinates in n-dimensional space. In a one-dimensional matrix (i.e. a number), there is only one coordinate, and the determinant can be interpreted as the (signed) length of a vector from the origin to this point. For a 2×2 matrix we have two coordinates in a plane, and the determinant is the (signed) area of the parallelogram that includes these two points and the origin. For a 3×3 matrix the determinant is the (signed) volume of the parallelepiped that includes the three points (in three-dimensional space) defined by the matrix and the origin. This is illustrated below. The idea extends up to higher dimensions in a similar way. In some sense then, the determinant is therefore related to the size of a matrix.



The algebraic definition of the determinant of an nxn matrix is a sum over all the possible products (permutations) of n elements taken from different rows and columns. The number of terms in the sum is $n!$, the number of possible permutations of n values (i.e. 2 for a 2×2 matrix, 6 for a 3×3 matrix etc). Each term in the sum is given a positive or a negative sign depending on whether the number of *permutation inversions* in the product is even or odd. A permutation inversion is just a pair of elements that are out of order when described by their indices. For example, for a set of four elements (a_1, a_2, a_3, a_4), the permutation $a_1a_2a_3a_4$ has all the elements in their correct order (i.e. in order of increasing index). However, the permutation $a_2a_4a_1a_3$ contains the permutation inversions a_2a_1, a_4a_1, a_4a_3.

For example, for a two-dimensional matrix

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{F.14}$$

where the subscripts label the row and column positions of the elements, there are 2 possible products/permutations involving elements from different rows and column, $a_{11}a_{22}$ and $a_{12}a_{21}$. In the second term, there is a permutation inversion involving the column indices 2 and 1 (permutation inversions involving the row and column indices should be looked for separately) so this term takes a negative sign, and the determinant is $a_{11}a_{22}$ - $a_{12}a_{21}$.

For a 3×3 matrix

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \tag{F.15}$$

the possible combinations of elements from different rows and columns, together with the sign from the number of permutations required to put their indices in numerical order are:

$$\begin{aligned} &a_{11}a_{22}a_{33} && (0 \text{ inversions}) \\ -&a_{11}a_{23}a_{32} && (1 \text{ inversion - } 3 > 2 \text{ in the column indices}) \\ -&a_{12}a_{21}a_{33} && (1 \text{ inversion - } 2 > 1 \text{ in the column indices}) \\ &a_{12}a_{23}a_{31} && (2 \text{ inversions - } 2 > 1 \text{ and } 3 > 1 \text{ in the column indices}) \\ &a_{13}a_{21}a_{32} && (2 \text{ inversions - } 3 > 1 \text{ and } 3 > 2 \text{ in the column indices}) \\ -&a_{13}a_{22}a_{31} && (3 \text{ inversions - } 3 > 2, 3 > 1, \text{ and } 2 > 1 \text{ in the column indices}) \end{aligned} \tag{F.16}$$

and the determinant is simply the sum of these terms.

This may all seem a little complicated, but in practice there is a fairly systematic procedure for calculating determinants. The determinant of a matrix A is usually written $\det(A)$ or $|a|$.

For a 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; det(A) = |A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \tag{F.17}$$

For a 3×3 matrix

$$B = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}; det(B) = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix} \tag{F.18}$$

For a $4x4$ matrix

$$C = \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{pmatrix}; \tag{F.19}$$

$$det(C) = a\begin{vmatrix} f & g & h \\ j & k & l \\ n & o & p \end{vmatrix} - b\begin{vmatrix} e & g & h \\ i & k & l \\ m & o & p \end{vmatrix} + c\begin{vmatrix} e & f & h \\ i & j & l \\ m & n & p \end{vmatrix} - d\begin{vmatrix} e & f & g \\ i & j & k \\ m & n & o \end{vmatrix}$$

and so on in higher dimensions. Note that the submatrices in the 3×3 example above are just the matrices formed from the original matrix B that don't include any elements from the same row or column as the premultiplying factors from the first row.

Matrix determinants have a number of important properties:

i. The determinant of the identity matrix is 1.

$$e.g. \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \tag{F.20}$$

ii. The determinant of a matrix is the same as the determinant of its transpose i.e. det(a) = det(A^T)

$$e.g. \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} \tag{F.21}$$

iii. The determinant changes sign when any two rows or any two columns are interchanged

$$e.g. \begin{vmatrix} a & b \\ c & d \end{vmatrix} = -\begin{vmatrix} b & a \\ d & c \end{vmatrix} = -\begin{vmatrix} c & d \\ a & b \end{vmatrix} = \begin{vmatrix} d & c \\ b & a \end{vmatrix} \tag{F.22}$$

iv. The determinant is zero if any row or column is entirely zero, or if any two rows or columns are equal or a multiple of one another.

$$e.g. \begin{vmatrix} 1 & 2 \\ 0 & 0 \end{vmatrix} = 0, \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0 \tag{F.23}$$

v. The determinant is unchanged by adding any linear combination of rows (or columns) to another row (or column).

vi. The determinant of the product of two matrices is the same as the product of the determinants of the two matrices i.e. det(AB) = det(A)det(B).

The requirement that in order for a matrix to have an inverse it must have a non-zero determinant follows from property vi). As mentioned previously, the product of a matrix and its inverse yields the identity matrix I. We therefore have:

$$\begin{aligned} det(A^{-1}A) = det(A^{-1})det(A) &= det(I) \\ det(A^{-1}) &= det(I)/det(A) = 1/det(A) \end{aligned} \tag{F.24}$$

It follows that a matrix A can only have an inverse if its determinant is non-zero, otherwise the determinant of its inverse would be undefined.

F: Matrices is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.8: Review of Matrices** by Claire Vallance is licensed CC BY 4.0. Original source: http://vallance.chem.ox.ac.uk/pdfs/SymmetryLectureNotes.pdf.

https://chem.libretexts.org/@go/page/198736

G: Numerical Methods

G: Numerical Methods is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

H: Partial Differentiation

The development of thermodynamics would have been unthinkable without calculus in more than one dimension (multivariate calculus) and partial differentiation is essential to the theory.

'Active' Variables

When applying partial differentiation it is very important to keep in mind, which symbol is the variable and which ones are the constants. Mathematicians usually write the variable as x or y and the constants as a, b or c but in Physical Chemistry the symbols are different. It sometimes helps to replace the symbols in your mind.

For example the van der Waals equation can be written as:

$$P = \frac{RT}{\overline{V} - b} - \frac{a}{\overline{V}^2} \tag{H.1}$$

Suppose we must compute the partial differential

$$\left(\frac{\partial P}{\partial \overline{V}}\right)_T$$

In this case molar volume is the variable 'x' and the pressure is the function $f(x)$, the rest is just constants, so Equation H.1 can be rewritten in the form

$$f(x) = \frac{c}{x - b} - \frac{a}{x^2} \tag{H.2}$$

When calculating

$$\left(\frac{\partial P}{\partial T}\right)_{\overline{V}}$$

should look at Equation H.1 as:

$$f(x) = cx - d$$

The active variable 'x' is now the temperature T and all the rest is just constants. It is useful to train your eye to pick out the one active one from all the inactive ones. Use highlighters, underline, rewrite, do whatever helps you best.

Cross Derivatives

As shown in Equations H.5 and H.6 there are also higher order partial derivatives versus T and versus V. A very interesting derivative of second order and one that is used extensively in thermodynamics is the mixed second order derivative.

$$\left(\frac{\partial^2 P}{\partial T\, \partial \overline{V}}\right) = \left(\frac{\partial^P}{\partial \overline{V}\, \partial T}\right) \tag{H.3}$$

Of course here the 'active' variable is first T, then V. The interesting thing about it is that it does not matter whether you first take T and then V or the other way around.

Example H-2 shows an example of how mixed derivatives can be used to translate one quantity into the other. This trick is used over and over again in thermodynamics because it allows you to replace a quantity that is really hard to measure by one (or more) that are much easier to get good experimental values for.

For example:

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial P}{\partial T}\right)_V$$

This expression is not obvious at all. It tells you that if you study the pressure P when heating up while keeping the volume the same (which is doable) you're measuring how the entropy changes with volume under isothermal conditions. Entropy will be discussed later, suffice it to say that nobody has ever constructed a working 'entropometer'! So that is an impossible quantity to measure directly.

The Decomposition of Changes

A very important result of multivariate calculus is that if a quantity Q is a function of more than one variable, say A and B that we can decompose any infinitesimal change dQ into infinitesimal changes in A and B in a very simple linear way:

$$dQ = \alpha\, dA + \beta dB \tag{H.4}$$

dq is sometimes referred to as the total differential. The coefficients α and β are the partial derivatives of first order versus A and B. This mathematical fact is something we will be using over and over.

Exact and Inexact differentials: State and path functions

The car trip

Suppose you drive your car up and down a mountain. You perform two measurements: you have a barometer that measures the air pressure and you keep an eye on your gas gage. Even though the barometer will show lower values on top of the mountain, its value will return to its initial value when you return home (barring weather changes). You might wish the same would hold for your gas gage particularly at current gas prices!

Pressure is a good example of a **state function** (it returns to its old value if you go back to a previous state). The other (the gas gage) is a path function. (Make a detour and your bank account will tell you difference!).

The difference between state and path functions has its roots deep in mathematics and it comes in as soon as a function has two of more variables.

The gas law is a good example. The pressure depends on both temperature T and (molar) volume V. When changing the pressure a little bit, say by dP we can show that we can write that out in the two possible components dT and dV as:

$$dP = pdT + qdV \tag{H.5}$$

$$= \left(\frac{\partial S}{\partial V}\right)_V dT + \left(\frac{\partial P}{\partial V}\right)_T dV \tag{H.6}$$

At first, I wrote arbitrary coefficients p and q in Equation H.5, but as you can see they are really partial derivatives (Equation H.6). This is another way that thermodynamics exploits multivariate calculus: it shows how total changes can be built up of various contributions.

The interesting thing is that ***if*** the function P is a ***state*** function (and your barometer will testify to that) then Equation H.3 **must** hold. However, if the function is a ***path*** function, then this equality does ***not*** hold.

Thermodynamics is largely based upon exploiting the above facts:

- It tries to define state functions to describe energy changes
- It tries to decompose changes into well-defined contributions
- It uses partial differentials to link known quantities to unknown ones

H: Partial Differentiation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **32.3: Partial Differentiation** is licensed CC BY-NC-SA 4.0.

I: Series and Limits

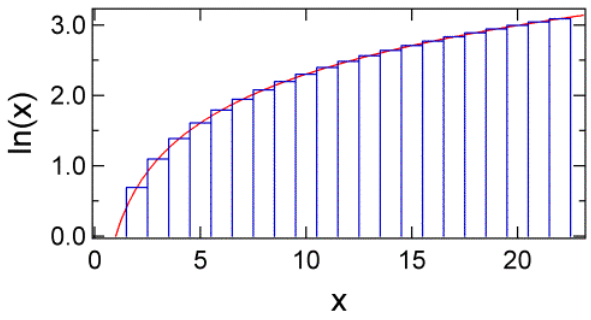
I: Series and Limits is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

J: The Binomial Distribution and Stirling's Appromixation

Stirling's approximation is named after the Scottish mathematician James Stirling (1692-1770). In confronting statistical problems we often encounter factorials of very large numbers. The factorial $N!$ is a product $N(N-1)(N-2)...(2)(1)$. Therefore, $\ln N!$ is a sum

$$\ln N! = \ln 1 + \ln 2 + \ln 3 + ... + \ln N = \sum_{k=1}^{N} \ln k. \tag{J.1}$$

where we have used the property of logarithms that $\log(abc) = \log(a) + \log(b) + \log(c)$. The sum is shown in figure below.



Using **Euler-MacLaurin** formula one has

$$\sum_{k=1}^{N} \ln k = \int_{1}^{N} \ln x\, dx + \sum_{k=1}^{p} \frac{B_{2k}}{2k(2k-1)} \left(\frac{1}{n^{2k-1}} - 1 \right) + R, \tag{J.2}$$

where $B_1 = -1/2$, $B_2 = 1/6$, $B_3 = 0$, $B_4 = -1/30$, $B_5 = 0$, $B_6 = 1/42$, $B_7 = 0$, $B_8 = -1/30$, ... are the **Bernoulli numbers**, and \(R\) is an error term which is normally small for suitable values of \(p\).

Then, for large N,

$$\ln N! \sim \int_{1}^{N} \ln x\, dx \approx N \ln N - N. \tag{J.3}$$

after some further manipulation one arrives at (apparently Stirling's contribution was the prefactor of $\sqrt{2\pi}$)

$$N! = \sqrt{2\pi N}\, N^N e^{-N} e^{\lambda_N} \tag{J.4}$$

where

$$\frac{1}{12N+1} < \lambda_N < \frac{1}{12N}. \tag{J.5}$$

The sum of the area under the blue rectangles shown below up to N is $\ln N!$. As you can see the rectangles begin to closely approximate the red curve as m gets larger. The area under the curve is given the integral of $\ln x$.

$$\ln N! = \sum_{m=1}^{N} \ln m \approx \int_{1}^{N} \ln x\, dx \tag{J.6}$$

To solve the integral use integration by parts

$$\int u\, dv = uv - \int v\, dy \tag{J.7}$$

Here we let $u = \ln x$ and $dv = dx$. Then $v = x$ and $du = \frac{dx}{x}$.

$$\int_{0}^{N} \ln x\, dx = x \ln x \Big|_{0}^{N} - \int_{0}^{N} x \frac{dx}{x} \tag{J.8}$$

Notice that $x/x = 1$ in the last integral and $x \ln x$ is 0 when evaluated at zero, so we have

$$\int_0^N \ln x\, dx = N \ln N - \int_0^N dx \tag{J.9}$$

Which gives us Stirling's approximation: $\ln N! = N \ln N – N$. As is clear from the figure above Stirling's approximation gets better as the number N gets larger (Table $J.1$).

Table $J.1$: Evaluation of Approximation with absolute values

N	N!	ln N!	N ln N – N	Error
10	3.63×10^{6}	15.1	13.02	13.8%
50	3.04×10^{64}	148.4	145.6	1.88%
100	9.33×10^{157}	363.7	360.5	0.88%
150	5.71×10^{262}	605.0	601.6	0.56%

Calculators often overheat at 200!, which is all right since clearly result are converging. In thermodynamics, we are often dealing very large N (i.e., of the order of Avagadro's number) and for these values Stirling's approximation is excellent.

References

1. J. Stirling "Methodus differentialis, sive tractatus de summation et interpolation serierum infinitarium", London (1730). English translation by J. Holliday "The Differential Method: A Treatise of the Summation and Interpolation of Infinite Series" (1749)

J: The Binomial Distribution and Stirling's Appromixation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Stirling's Approximation** by SklogWiki, Stefan Franzen is licensed CC BY-NC-SA 4.0. Original source: http://www.sklogwiki.org/SklogWiki/index.php/Stirling%27s_approximation.

Index

A

Anharmonicity
- 5.3: The Harmonic Oscillator is an Approximation

Argand diagram
- A: Complex Numbers

B

Balmer series
- 1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum

beating
- 3.4: The Quantum Mechanical Free Particle

blackbody radiation
- 1.1: Blackbody Radiation Cannot Be Explained Classically

bond order
- 9.8: Molecular-Orbital Theory Does not Predict a Stable Diatomic Helium Molecule

C

Commuting Operators
- 4.6: Heisenburg Uncertainy Principle III - Commuting Operators

correspondence principle
- 3.3: The Schrödinger Equation is an Eigenvalue Problem

cyclic boundary condition
- 5.8: The Energy Levels of a Rigid Rotor

D

de Broglie wavelength
- 1.6: Matter Has Wavelike Properties

degeneracy
- 3.11: A Particle in a Three-Dimensional Box

Determinants
- E: Determinants

Diatomic Vibration
- 5.5: The Harmonic Oscillator and Infrared Spectra

E

Eigenfunction
- 3.3: The Schrödinger Equation is an Eigenvalue Problem

eigenvalue problem
- 3.3: The Schrödinger Equation is an Eigenvalue Problem
- 4.3: Observable Quantities Must Be Eigenvalues of Quantum Mechanical Operators

envelope function
- 3.4: The Quantum Mechanical Free Particle

exact differential
- H: Partial Differentiation

expectation value
- 3.8: The Average Momentum of a Particle in a Box is Zero

F

free particle
- 3.4: The Quantum Mechanical Free Particle

G

Gaussian Wavepacket
- 3.4: The Quantum Mechanical Free Particle

H

Hamiltonian
- 4.2: Quantum Operators Represent Classical Variables

Harmonic oscillator
- 5.1: A Harmonic Oscillator Obeys Hooke's Law
- 5.3: The Harmonic Oscillator is an Approximation
- 5.4: The Harmonic Oscillator Energy Levels
- 5.5: The Harmonic Oscillator and Infrared Spectra

Heisenberg uncertainty principle
- 3.9: The Uncertainty Principle Redux - Estimating Uncertainties from Wavefunctions

Helium Dimer
- 9.8: Molecular-Orbital Theory Does not Predict a Stable Diatomic Helium Molecule

Hermite polynomials
- 5.6: The Harmonic-Oscillator Wavefunctions Involve Hermite Polynomials
- 5.7: Hermite Polynomials are either Even or Odd Functions

Hermitian operators
- 4.5: The Eigenfunctions of Operators are Orthogonal

Hooke's Law
- 5.1: A Harmonic Oscillator Obeys Hooke's Law

Hydrogen Atomic Spectrum
- 1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum

I

inexact differential
- H: Partial Differentiation

IR spectra
- 5.5: The Harmonic Oscillator and Infrared Spectra

L

laplacian
- 4.2: Quantum Operators Represent Classical Variables

LCAO
- 9.4: Chemical Bond Stability

Lyman series
- 1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum

M

microwave spectroscopy
- 5.9: The Rigid Rotator is a Model for a Rotating Diatomic Molecule

Morse potential
- 5.3: The Harmonic Oscillator is an Approximation
- 5.5: The Harmonic Oscillator and Infrared Spectra

N

nodes
- 3.8: The Average Momentum of a Particle in a Box is Zero

Nonstationary States
- 4.4: The Time-Dependent Schrödinger Equation

O

Operators
- 3.2: Linear Operators in Quantum Mechanics
- 4.2: Quantum Operators Represent Classical Variables

Orthonormality
- 3.6: The Energy of a Particle in a Box is Quantized

overlap
- 9.3: The Overlap Integral

Overlap integral
- 9.3: The Overlap Integral
- 9.3B: Evaluating the Overlap Integral (Optional)

P

paramagnetic
- 9.10: Molecular Orbital Theory Predicts that Molecular Oxygen is Paramagnetic

Particle in a 3D Box
- 3.11: A Particle in a Three-Dimensional Box

particle in a box
- 3.6: The Energy of a Particle in a Box is Quantized

Pashen series
- 1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum

perturbation theory
- 7.4: Perturbation Theory Expresses the Solutions in Terms of Solved Problems

Perturbation Theory for Helium
- 8.2: Perturbation Theory and the Variational Method for Helium

Pfund series
- 1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum

Principle of Superposition
- 2.5: A Vibrating Membrane
- 3.4: The Quantum Mechanical Free Particle

Probabilistic Interpretation
- 3.5: Wavefunctions Have a Probabilistic Interpretation

R

reduced mass
- 5.2: The Equation for a Harmonic-Oscillator Model of a Diatomic Molecule Contains the Reduced Mass of the Molecule

reducing agent
- 5.5: The Harmonic Oscillator and Infrared Spectra

Rigid rotor
- 5.8: The Energy Levels of a Rigid Rotor
- 5.9: The Rigid Rotator is a Model for a Rotating Diatomic Molecule

Rydberg formula
- 1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum

S

Schmidt orthogonalization theorem
- 4.5: The Eigenfunctions of Operators are Orthogonal

Schrödinger equation
3.1: The Schrödinger Equation
Selection rules
5.5: The Harmonic Oscillator and Infrared Spectra
Separation of Variables
2.2: The Method of Separation of Variables
stationary state
4.4: The Time-Dependent Schrödinger Equation

T

Total Differentials
H: Partial Differentiation
Transition Energies
5.9: The Rigid Rotator is a Model for a Rotating Diatomic Molecule

U

Ultraviolet Catastrophe
1.1: Blackbody Radiation Cannot Be Explained Classically

V

Variational Method for Helium
8.2: Perturbation Theory and the Variational Method for Helium

W

Wavefunction Normalization
3.7: Wavefunctions Must Be Normalized
wavepackets
3.4: The Quantum Mechanical Free Particle
Wien's Displacement Law
1.1: Blackbody Radiation Cannot Be Explained Classically

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: CH-331 Text

Webpages: 114

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC-SA 4.0: 93.9% (107 pages)
- Undeclared: 6.1% (7 pages)

By Page

- CH-331 Text — *CC BY-NC-SA 4.0*
  - Front Matter — *Undeclared*
    - TitlePage — *Undeclared*
    - InfoPage — *Undeclared*
    - Table of Contents — *Undeclared*
  - 1: The Dawn of the Quantum Theory — *CC BY-NC-SA 4.0*
    - 1.1: Blackbody Radiation Cannot Be Explained Classically — *CC BY-NC-SA 4.0*
    - 1.2: Quantum Hypothesis used for Blackbody Radiation Law — *CC BY-NC-SA 4.0*
    - 1.3: Photoelectric Effect Explained with Quantum Hypothesis — *CC BY-NC-SA 4.0*
    - 1.4: The Hydrogen Atomic Spectrum — *CC BY-NC-SA 4.0*
    - 1.5: The Rydberg Formula and the Hydrogen Atomic Spectrum — *CC BY-NC-SA 4.0*
    - 1.6: Matter Has Wavelike Properties — *CC BY-NC-SA 4.0*
    - 1.7: de Broglie Waves can be Experimentally Observed — *CC BY-NC-SA 4.0*
    - 1.8: The Bohr Theory of the Hydrogen Atom — *CC BY-NC-SA 4.0*
    - 1.9: The Heisenberg Uncertainty Principle — *CC BY-NC-SA 4.0*
    - 1.E: The Dawn of the Quantum Theory (Exercises) — *CC BY-NC-SA 4.0*
  - 2: The Classical Wave Equation — *CC BY-NC-SA 4.0*
    - 2.1: The One-Dimensional Wave Equation — *CC BY-NC-SA 4.0*
    - 2.2: The Method of Separation of Variables — *CC BY-NC-SA 4.0*
    - 2.3: Oscillatory Solutions to Differential Equations — *CC BY-NC-SA 4.0*
    - 2.4: The General Solution is a Superposition of Normal Modes — *CC BY-NC-SA 4.0*
    - 2.5: A Vibrating Membrane — *CC BY-NC-SA 4.0*
    - 2.E: The Classical Wave Equation (Exercises) — *CC BY-NC-SA 4.0*
  - 3: The Schrödinger Equation & a Particle in a Box — *CC BY-NC-SA 4.0*
    - 3.1: The Schrödinger Equation — *CC BY-NC-SA 4.0*
    - 3.2: Linear Operators in Quantum Mechanics — *CC BY-NC-SA 4.0*
    - 3.3: The Schrödinger Equation is an Eigenvalue Problem — *CC BY-NC-SA 4.0*
    - 3.4: The Quantum Mechanical Free Particle — *CC BY-NC-SA 4.0*
    - 3.5: Wavefunctions Have a Probabilistic Interpretation — *CC BY-NC-SA 4.0*
    - 3.6: The Energy of a Particle in a Box is Quantized — *CC BY-NC-SA 4.0*
    - 3.7: Wavefunctions Must Be Normalized — *CC BY-NC-SA 4.0*
    - 3.8: The Average Momentum of a Particle in a Box is Zero — *CC BY-NC-SA 4.0*
    - 3.9B: Particle in a Finite Box and Tunneling (optional) — *CC BY-NC-SA 4.0*
    - 3.9: The Uncertainty Principle Redux - Estimating Uncertainties from Wavefunctions — *CC BY-NC-SA 4.0*
    - 3.10: A Particle in a Two-Dimensional Box — *CC BY-NC-SA 4.0*
    - 3.11: A Particle in a Three-Dimensional Box — *CC BY-NC-SA 4.0*
    - 3.E: The Schrödinger Equation and a Particle in a Box (Exercises) — *CC BY-NC-SA 4.0*
  - 4: Postulates and Principles of Quantum Mechanics — *CC BY-NC-SA 4.0*

- 4.1: The Wavefunction Specifies the State of a System — *CC BY-NC-SA 4.0*
- 4.2: Quantum Operators Represent Classical Variables — *CC BY-NC-SA 4.0*
- 4.3: Observable Quantities Must Be Eigenvalues of Quantum Mechanical Operators — *CC BY-NC-SA 4.0*
- 4.4: The Time-Dependent Schrödinger Equation — *CC BY-NC-SA 4.0*
- 4.5: The Eigenfunctions of Operators are Orthogonal — *CC BY-NC-SA 4.0*
- 4.6: Heisenburg Uncertainy Principle III - Commuting Operators — *CC BY-NC-SA 4.0*
- 4.E: Postulates and Principles of Quantum Mechanics (Exercises) — *CC BY-NC-SA 4.0*

5: The Harmonic Oscillator and the Rigid Rotor — *CC BY-NC-SA 4.0*

- 5.1: A Harmonic Oscillator Obeys Hooke's Law — *CC BY-NC-SA 4.0*
- 5.2: The Equation for a Harmonic-Oscillator Model of a Diatomic Molecule Contains the Reduced Mass of the Molecule — *CC BY-NC-SA 4.0*
- 5.3: The Harmonic Oscillator is an Approximation — *CC BY-NC-SA 4.0*
- 5.4: The Harmonic Oscillator Energy Levels — *CC BY-NC-SA 4.0*
- 5.5: The Harmonic Oscillator and Infrared Spectra — *CC BY-NC-SA 4.0*
- 5.6: The Harmonic-Oscillator Wavefunctions Involve Hermite Polynomials — *CC BY-NC-SA 4.0*
- 5.7: Hermite Polynomials are either Even or Odd Functions — *CC BY-NC-SA 4.0*
- 5.8: The Energy Levels of a Rigid Rotor — *CC BY-NC-SA 4.0*
- 5.9: The Rigid Rotator is a Model for a Rotating Diatomic Molecule — *CC BY-NC-SA 4.0*
- 5.E: The Harmonic Oscillator and the Rigid Rotor (Exercises) — *CC BY-NC-SA 4.0*

6: The Hydrogen Atom — *CC BY-NC-SA 4.0*

- 6.1: The Schrödinger Equation for the Hydrogen Atom Can Be Solved Exactly — *CC BY-NC-SA 4.0*
- 6.2: The Wavefunctions of a Rigid Rotator are Called Spherical Harmonics — *CC BY-NC-SA 4.0*
- 6.3: The Three Components of Angular Momentum Cannot be Measured Simultaneously with Arbitrary Precision — *CC BY-NC-SA 4.0*
- 6.4: Hydrogen Atomic Orbitals Depend upon Three Quantum Numbers — *CC BY-NC-SA 4.0*
- 6.5: s Orbitals are Spherically Symmetric — *CC BY-NC-SA 4.0*
- 6.6: Orbital Angular Momentum and the p-orbitals — *CC BY-NC-SA 4.0*
- 6.7: The Helium Atom Cannot Be Solved Exactly — *CC BY-NC-SA 4.0*
- 6.E: The Hydrogen Atom (Exercises) — *CC BY-NC-SA 4.0*

7: Approximation Methods — *CC BY-NC-SA 4.0*

- 7.0: Slater's Rules for Shielding — *CC BY-NC-SA 4.0*
- 7.1: The Variational Method — *CC BY-NC-SA 4.0*
- 7.2: Linear Variational Method and the Secular Determinant — *CC BY-NC-SA 4.0*
- 7.3: Trial Functions Can Be Linear Combinations of Functions with Variational Parameters — *CC BY-NC-SA 4.0*
- 7.4: Perturbation Theory Expresses the Solutions in Terms of Solved Problems — *CC BY-NC-SA 4.0*
- 7.E: Approximation Methods (Exercises) — *CC BY-NC-SA 4.0*

8: Multielectron Atoms — *CC BY-NC-SA 4.0*

- 8.1: Atomic and Molecular Calculations are Expressed in Atomic Units — *CC BY-NC-SA 4.0*
- 8.2: Perturbation Theory and the Variational Method for Helium — *CC BY-NC-SA 4.0*
- 8.3: Hartree-Fock Equations are Solved by the Self-Consistent Field Method — *CC BY-NC-SA 4.0*
- 8.4: An Electron Has An Intrinsic Spin Angular Momentum — *CC BY-NC-SA 4.0*
- 8.5: Wavefunctions must be Antisymmetric to Interchange of any Two Electrons — *CC BY-NC-SA 4.0*
- 8.6: Antisymmetric Wave Functions can be Represented by Slater Determinants — *CC BY-NC-SA 4.0*
- 8.7: Hartree-Fock Calculations Give Good Agreement with Experimental Data — *CC BY-NC-SA 4.0*
- 8.8: Term Symbols Gives Detailed Descriptions of an Electron Configuration — *CC BY-NC-SA 4.0*
- 8.8B: Multi-electron Considerations - A Closer Look at Helium — *CC BY-NC-SA 4.0*
- 8.9: The Allowed Values of J - the Total Angular Momentum Quantum Number — *CC BY-NC-SA 4.0*
- 8.10: Hund's Rules Determine the Term Symbols of the Ground Electronic States — *CC BY-NC-SA 4.0*
- 8.11: Using Atomic Term Symbols to Describe Atomic Spectra — *CC BY-NC-SA 4.0*
- 8.E: Multielectron Atoms (Exercises) — *CC BY-NC-SA 4.0*

9: The Chemical Bond: Diatomic Molecules — *CC BY-NC-SA 4.0*

- 9.1: The Born-Oppenheimer Approximation Simplifies the Molecular Schrödinger Equation — *CC BY-NC-SA 4.0*
- 9.2: The H_2^+ Prototypical Species — *CC BY-NC-SA 4.0*
- 9.2B: Solving the H_2^+ System Exactly (Optional) — *CC BY-NC-SA 4.0*
- 9.3: The Overlap Integral — *CC BY-NC-SA 4.0*
- 9.3B: Evaluating the Overlap Integral (Optional) — *CC BY-NC-SA 4.0*
- 9.4: Chemical Bond Stability — *CC BY-NC-SA 4.0*
- 9.5: Bonding and Antibonding Orbitals — *CC BY-NC-SA 4.0*
- 9.6: A Simple Molecular-Orbital Treatment of H_2 Places Both Electrons in a Bonding Orbital — *CC BY-NC-SA 4.0*
- 9.7: Molecular Orbitals Can Be Ordered According to Their Energies — *CC BY-NC-SA 4.0*
- 9.8: Molecular-Orbital Theory Does not Predict a Stable Diatomic Helium Molecule — *CC BY-NC-SA 4.0*
- 9.9: Electrons Populate Molecular Orbitals According to the Pauli Exclusion Principle — *CC BY-NC-SA 4.0*
- 9.10: Molecular Orbital Theory Predicts that Molecular Oxygen is Paramagnetic — *CC BY-NC-SA 4.0*
- 9.E: The Chemical Bond: Diatomic Molecules (Exercises) — *CC BY-NC-SA 4.0*

MathChapters — *CC BY-NC-SA 4.0*

- A: Complex Numbers — *CC BY-NC-SA 4.0*
- B: Probability and Statistics — *CC BY-NC-SA 4.0*
- C: Vectors — *CC BY-NC-SA 4.0*
- D: Spherical Coordinates — *CC BY-NC-SA 4.0*
- E: Determinants — *CC BY-NC-SA 4.0*
- F: Matrices — *CC BY-NC-SA 4.0*
- G: Numerical Methods — *CC BY-NC-SA 4.0*
- H: Partial Differentiation — *CC BY-NC-SA 4.0*
- I: Series and Limits — *CC BY-NC-SA 4.0*
- J: The Binomial Distribution and Stirling's Appromixation — *CC BY-NC-SA 4.0*

Back Matter — *Undeclared*

- Index — *Undeclared*
- Glossary — *Undeclared*