José J.C. Teixeira-Dias

Molecular Physical Chemistry

A Computer-based Approach using Mathematica® and Gaussian

Springer

Molecular Physical Chemistry

José J.C. Teixeira-Dias

Molecular Physical Chemistry

A Computer-based Approach using Mathematica® and Gaussian

Springer

José J.C. Teixeira-Dias
University of Aveiro
Aveiro, Baixo Vouga
Portugal

ISBN 978-3-319-41092-0 ISBN 978-3-319-41093-7 (eBook)
DOI 10.1007/978-3-319-41093-7

Library of Congress Control Number: 2016947376

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book contains chapters on thermodynamics, chemical kinetics, quantum chemistry, molecular symmetry, molecular structure, crystals, and water, and is intended for second-year master's students in chemistry. It presents the subject through real examples, discussing the results of molecular orbital calculations performed by *Gaussian* on small molecules, exploring and running *Mathematica* codes presented at the end of each chapter that enable the student to plot functions, normalize functions, fit data, solve equations, and test physical models; they are accompanied by detailed explanations that provide insight and a suitable environment for active learning. Each chapter contains a glossary of important scientific and technical terms, and the book includes detailed and complete answers to all exercises. Since the molecular orbital calculations presented are standard, packages other than *Gaussian* can alternatively be used to provide the necessary data. Students who are unfamiliar with *Mathematica* should watch the set of short videos provided by this software to learn to write and run small programs and follow the explanations to the selected codes at the end of each chapter. Those who are familiar with other computational tools can alternatively use them.

Aveiro, Portugal — José J.C. Teixeira-Dias

Contents

Credits

Gaussian **(www.gaussian.com)**
The computational quantum chemistry calculations whose results are presented in this book were carried out with the *Gaussian* 09 series of programs.

Authors for *Gaussian* 09, Revision A.02: M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery, Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, and D.J. Fox, Gaussian, Inc., Wallingford CT, 2009.

GaussView **(www.gaussian.com)**
GaussView is the graphical interface to *Gaussian*.

Authors for *GaussView*, Version 5.0.9: Roy Dennington, Todd Keith, and John Millam, Semichem Inc., Shawnee Mission, KS, 2009.

Natural bond orbitals
A natural bond orbitals analysis of the electron population is included in *Gaussian* 09.

Reference: NBO Version 3.1, E.D. Glendening, A.E. Reed, J.E. Carpenter, and F. Weinhold.

Mathematica **(www.wolfram.com)**
Author of *Mathematica*: Wolfram Research Inc. Champaign, IL, 2010.

Wolfram Alpha **(www.wolframalpha.com)**
This is a computational knowledge-based engine that provides answers to questions raised through the Internet in free-form input of natural language (English).

Wolfram MathWorld (mathworld.wolfram.com)
This Internet site, built with *Mathematica* technology, is considered to be the web's most extensive mathematical resource.

CCCBDB (cccbdb.nist.gov)
National Institute of Standards and Technology (NIST) Computational Chemistry Comparison and Benchmark DataBase

NIST Standard Reference Database Number 101

Release 17b, September 2015, Editor: Russell D. Johson III

This database consists of experimental and computational thermochemical data for a selected set of more than 1700 gas phase atoms and small molecules.

Character tables (**symmetry.jacobs-university.de**)
These character tables apply to chemically important point groups.

Handbook of Chemistry and Physics, 2011
On CD-ROM, Version 2011, W.M. Haynes, Editor-in-Chief, CRC Press.

Thermodynamics

1

Abstract

The first three sections of this chapter include the ideal gas equation, the kinetic model of gases, and the Van der Waals equation. They aim at building a unified background for thermodynamics. Having in mind those students who might need some refreshment of the mathematical concepts much used in thermodynamics, such as exact and inexact differentials, the fundamental theorem of calculus, and line integrals, we introduce a short section on these mathematical tools, followed by sections on thermodynamic systems, heat and work, and the first law. In order to formally arrive at the state function entropy, we show that an arbitrary reversible cycle can be covered to the required precision by a mesh of infinitesimal Carnot cycles and arrive at the Clausius inequality for an arbitrary irreversible cycle. Following Prigogine, a distinction is then established between entropy transfer across system boundaries and entropy produced inside a system due to irreversible processes that drive the system to equilibrium. These are illustrated with heat flow, gas expansion, matter diffusion, and a chemical reaction in the gas phase. Next, we derive the Gibbs–Duhem equation and obtain the chemical potential for the ideal gas, real gases, liquid solutions, pure liquids, and solids, before presenting sections on the Gibbs energy, chemical equilibrium, the Gibbs phase rule, and the Helmholtz energy. The last three sections of this chapter, on surface tension, membrane potential, and the electrochemical cell, are important applications of the first and second laws of thermodynamics. At the end of the chapter we present several Mathematica codes (Maxwell Distribution of Molecular Speeds, Critical Point for Van der Waals Fluid, 3D Plot of the Van der Waals Equation in Reduced Variables, Absolute and Reduced Temperatures for the Van der Waals Carbon Dioxide, Isothermal and Adiabatic Transformations on an Ideal Gas Surface, Efficiency of the Carnot Heat Engine, Gibbs Energy, and Affinity of a Chemical Reaction) that are interconnected with the main text and contain detailed explanations of important

J.J.C. Teixeira-Dias, *Molecular Physical Chemistry*,
DOI 10.1007/978-3-319-41093-7_1

commands. Some of these explanations present suggestions for the student to follow. The student will also find at the end of this chapter a glossary of important scientific and technical terms in thermodynamics and a list of exercises, whose complete answers are at the end of the book.

1.1 Ideal Gas

In, 1660, Robert Boyle (1627–1691) showed that at a defined temperature T, the volume V of a gas is inversely proportional to its pressure p, that is,

$$V = f_1(T)/p \tag{1.1}$$

where $f_1(T)$ is a function of temperature (see Note **§1**). Edmé Mariotte [1620(?)–1684] independently confirmed this functional dependence between volume and pressure at constant temperature. In the last quarter of the eighteenth century, Jacques Charles (1746–1823) studied the dependence between volume and temperature, having found that the volume is proportional to the temperature at constant pressure,

$$V = Tf_2(p) \tag{1.2}$$

In 1811, Amedeo Avogadro (1776–1856) put forward a hypothesis according to which equal volumes of any two gases at the same temperature and pressure contain the same number of molecules. This hypothesis implies that at a defined temperature and pressure, the volume of a gas is proportional to the number of gas molecules N. Using the **Avogadro constant** ($N_A = 6.022 \times 10^{23}$ mol^{-1}, see Appendix), the amount of gas n is given by

$$n = N/N_A \tag{1.3}$$

Combining (1.1), (1.2) and the Avogadro hypothesis (gas volume proportional to the amount of gas), one concludes that $f_1(T)$ is given by

$$f_1(T) = \text{constant} \times nT = nRT \tag{1.4}$$

where R is called the **gas constant** ($R = 8.3144$ J mol^{-1} K^{-1}; see Appendix). Substitution of (1.4) in (1.1) leads to

$$pV = nRT \tag{1.5}$$

The gas constant is related to the **Boltzmann constant** (k_B = 1.3806 × 10^{-23} J K^{-1}; see Appendix) by

$$R = k_B N_A \tag{1.6}$$

A gas whose pressure p, volume V, temperature T, and amount of substance n obey (1.5) is called an **ideal gas**. Since these four variables are linked by (1.5), three of them determine the fourth. For instance, n, V, and T determine p. A gas with defined values of pressure, volume, and temperature is said to be in a **thermodynamic state**. For that reason, (1.5) is called the **equation of state**, and the variables p, V, T, and n are called **state variables**. Substitution of (1.6) and (1.3) in (1.5) leads to

$$p\,V = N\,k_B T \tag{1.7}$$

Dividing both members of (1.7) by V gives

$$p = \rho_N k_B T \tag{1.8}$$

where ρ_N is the **number density**, given by

$$\rho_N = N/V \tag{1.9}$$

Equations (1.5), (1.7), and (1.8) are equivalent forms of the ideal gas equation of state.

1.2 Kinetic Model of Gases

A gas can be studied at the macroscopic and molecular levels. The macroscopic study requires the measurement of variables such as pressure, volume, temperature, and amount of gas, interrelated by an empirical, relationship involving these variables (the equation of state), as we have just shown for the ideal gas. In turn, at the molecular level, the study of gases requires a theoretical model that allows for the understanding of the dynamical behavior of molecules under the model assumptions and the evaluation of macroscopic intensive variables such as pressure and temperature.

It was Daniel Bernoulli (1700–1782) who first concluded that the pressure of a gas results from collisions of gas particles with the container walls. At that time, no one knew about molecules or could predict the order of magnitude of the speed of gas particles.

The kinetic model of gases explains macroscopic gas properties such as pressure and temperature from the random motion of molecules and was mostly developed during the nineteenth century, with main contributions by James Prescott Joule

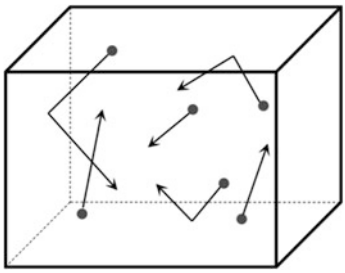


Fig. 1.1 Molecules of a gas are in constant random motion

(1818–1889), James Clerk Maxwell (1831–1879) and Ludwig Boltzmann (1844–1906). This gas model is based on five assumptions. The first states that the average distance between molecules in a gas is far greater than the dimensions of the molecules. The second says that the molecules of a gas are in constant random motion (Fig. 1.1). The third assumption sees each molecule as a sphere making elastic collisions with other molecules or with the container walls. The fourth assumption says that the molecules do not interact with each other except during collisions, and the last assumption states that gas molecules obey Newton's laws of motion.

At equilibrium, each gas molecule exists in an **isotropic space**; that is, all directions emerging from the molecular center of mass are equivalent, and the choice of directions for a reference system of axes is arbitrary. Therefore, equations can be deduced for the x velocity component of a molecule and easily generalized for the y and z components.

In the kinetic model of gases, molecules correspond to rigid spheres, that is, spheres that cannot be deformed. Therefore, the kinetic energy is not transferred into potential energy. Kinetic energy is the single form of energy existent in the kinetic model of gases. Before and after a collision between molecules, the total kinetic energy (sum of the kinetic energies of both molecules) is invariant. The same happens with the collision of a molecule with a container wall, but in this case, the kinetic energy of the container wall is zero, and so the colliding molecule has the same kinetic energy before and after the collision. In the kinetic model of gases, all the collisions are elastic.

1.2.1 Pressure and Temperature

Among the molecules that move toward the container wall at the right in Fig. 1.2, those that collide with the wall within the time interval Δt cannot be at a distance from the wall greater than $v_x \Delta t$, where $v_x > 0$. The volume occupied by such molecules is $v_x \Delta t A$, where A is the area of the wall, and the number of such molecules is given by

$$v_x \Delta t A \rho_N / 2 \tag{1.10}$$

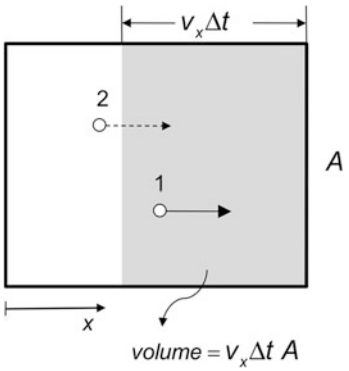


Fig. 1.2 In the time interval Δt, molecule 1 collides with the wall at the right

where ρ_N is the number density given by (1.9). The factor of 1/2 in (1.10) accounts for the fact that one-half of the molecules move toward each of the walls, right and left. In each collision with the wall at the right, the x component of the velocity changes sign. Hence, the variation of linear momentum on collision with the wall is given by $mv_x - (-mv_x) = 2mv_x$ (Fig. 1.3), and the total variation of linear momentum is $2mv_x$ times the number of molecules (1.10), that is,

$$mv_x^2 \Delta t A \rho_N \tag{1.11}$$

According to Newton's equation, the force is equal to the time derivative of the linear momentum, which is given approximately by the ratio of (1.11) to Δt,

$$mv_x^2 A \rho_N \tag{1.12}$$

since the time interval is very short. Therefore, the pressure (force per unit area) on the right wall is given by (1.12) divided by A, that is,

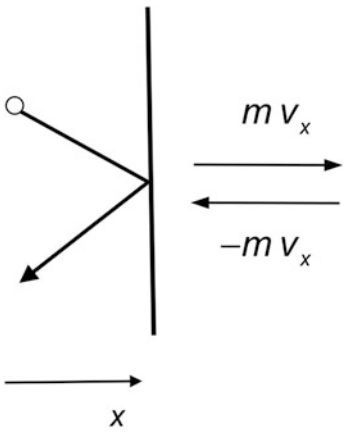


Fig. 1.3 Molecule in elastic collision with a wall

$$p = mv_x^2 \rho_N \tag{1.13}$$

The velocity axes are orthogonal, and so we can write

$$v^2 = v_x^2 + v_y^2 + v_z^2 \tag{1.14}$$

For an isotropic space,

$$v^2 = 3v_x^2 \tag{1.15}$$

Substitution of (1.15) in the second member of (1.13) leads to

$$p = mv^2 \rho_N / 3 \tag{1.16}$$

However, not all molecular velocities have the same absolute value (the same speed), and consequently the same square speed. In the ensemble average of square speeds, the average is taken over all N gas molecules in the container,

$$\langle v^2 \rangle = \left(v_1^2 + v_2^2 + \cdots + v_N^2\right)/N \tag{1.17}$$

and we can write

$$p = m\langle v^2 \rangle \rho_N / 3 \tag{1.18}$$

Substitution of (1.18) in (1.8) leads to

$$3k_B T/2 = m\langle v^2 \rangle /2 = \langle E_{cin} \rangle \tag{1.19}$$

This equation shows that temperature is proportional to the ensemble average of the kinetic energy of gas molecules.

1.2.2 Distribution of Velocities

In velocity space, the Cartesian axes correspond to the velocity components v_x, v_y, and v_z. The probability of finding the velocity vector $\boldsymbol{v} = (v_x, v_y, v_z)$ in the domain $(v_x, v_x + dv_x)$, $(v_y, v_y + dv_y)$, $(v_z, v_z + dv_z)$ (Fig. 1.4) is given by

$$P(\boldsymbol{v}) dv_x dv_y dv_z \tag{1.20}$$

where the function $P(\boldsymbol{v})$ is the probability density for the distribution of molecular velocities, first evaluated by James Clerk Maxwell for a gas in thermodynamic equilibrium and given by

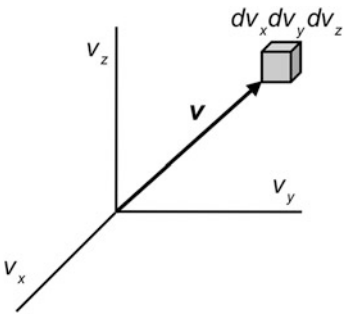


Fig. 1.4 Velocity space

$$P(\boldsymbol{v}) \propto \exp\left[-m\left(v_x^2 + v_y^2 + v_z^2\right)/(2k_{\mathrm{B}}T)\right] \tag{1.21}$$

where the symbol $\propto$ means that the proportionality constant of the above expression has been omitted and the absolute value of the exponent of (1.21) is the molecular kinetic energy $m(v_x^2 + v_y^2 + v_z^2)/2$ divided by $k_{\mathrm{B}}T$. The proportionality constant of (1.21) is given by

$$1/\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left[-m\left(v_x^2 + v_y^2 + v_z^2\right)/(2k_{\mathrm{B}}T)\right] \mathrm{d}v_x \mathrm{d}v_y \mathrm{d}v_z = \left(\frac{m}{2\pi k_{\mathrm{B}}T}\right)^{3/2} \tag{1.22}$$

(**E1**). Therefore, the probability density is

$$P(\boldsymbol{v}) = \left(\frac{m}{2\pi\, k_{\mathrm{B}}T}\right)^{3/2} \exp\left[-m\left(v_x^2 + v_y^2 + v_z^2\right)/(2k_{\mathrm{B}}T)\right] \tag{1.23}$$

and is said to be normalized, that is, the integral of the probability density over all velocity space is equal to 1. The density $P(\boldsymbol{v})$ is called the **Maxwell velocity distribution**. Expression (1.23) can be easily factored as the product of the v_x, v_y, and v_z components of the velocity vector $\boldsymbol{v}$, that is,

$$P(\boldsymbol{v}) = P(v_x)P(v_y)P(v_z) \tag{1.24}$$

where

$$P(v_x) = \left(\frac{m}{2\pi\, k_{\mathrm{B}}T}\right)^{1/2} \exp\left[-mv_x^2/(2k_{\mathrm{B}}T)\right] \tag{1.25}$$

Since $P(v_x)$ depends on v_x^2, both v_x and $-v_x$ give the same value to the function, which is symmetric about zero, that is, its mean value is zero. Figure 1.5 shows $P(v_x)$ for molecular oxygen, at 100 and 300 K.

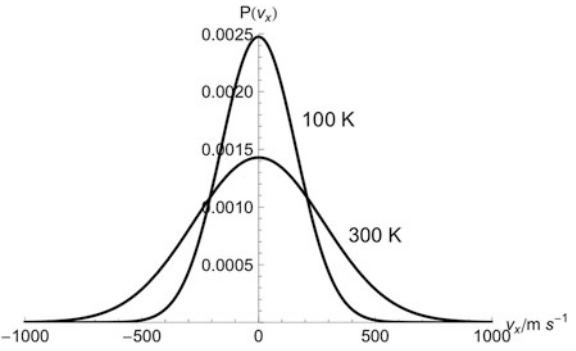


Fig. 1.5 $P(v_x)$ at 100 and 300 K, for molecular oxygen. Figure obtained with *Mathematica*

For an isotropic space,

$$P(v_x) = P(v_y) = P(v_z) \tag{1.26}$$

and the geometric place for the tips of velocity vectors with the same absolute value is the volume between two concentric spheres with radii v and $v + \mathrm{d}v$, which is given by the product of the area of the spherical surface $4\pi v^2$ and $\mathrm{d}v$. Thus, using (1.23) and (1.14), we can write

$$f(v)\mathrm{d}v = P(\boldsymbol{v})4\pi v^2 \mathrm{d}v = 4\pi\left(\frac{m}{2\pi k_\mathrm{B} T}\right)^{3/2} \exp\left[-mv^2/(2k_\mathrm{B}T)\right] v^2 \mathrm{d}v \tag{1.27}$$

where $f(v)$ is the **distribution of molecular speeds** [the word *speed* is used to express the absolute value of the *velocity* vector, and the integration of (1.27) is now between zero and infinity]. Figure 1.6 shows the distributions of molecular speeds for oxygen, at 100 and 300 K (see *Mathematica* code **M1**). It can be seen that these distributions are no longer symmetric, due to the factor v^2, which strongly weights larger values of molecular speeds, thus widening the higher wing of the distributions.

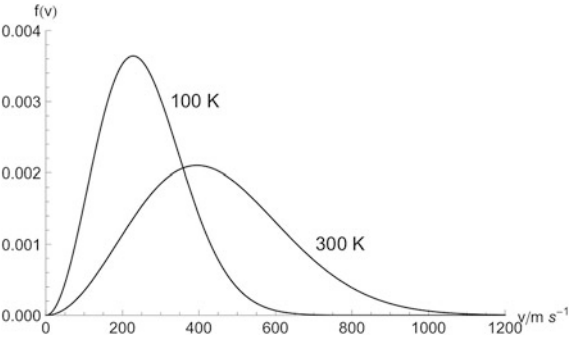


Fig. 1.6 Distributions of molecular speeds for oxygen, at 100 and 300 K. Figure obtained with *Mathematica*

At a defined temperature, the speed that corresponds to the maximum of $f(v)$ is called **the most probable speed**, v_p, and is given by the positive solution of the equation

$$\left(\frac{\partial f}{\partial v}\right)_T = 0 \tag{1.28}$$

[$f(v)$ has a single positive stationary point that is the curve maximum]. The positive solution for (1.28) is

$$v_p = \sqrt{\frac{2k_B T}{m}} \tag{1.29}$$

At 273.15 K, the most probable speed for O_2 is 377 m s^{-1}.

As mentioned before, $f(v)dv$ is the probability for finding the speed v. Therefore, $f(v)dv$ is the weighting factor for the speed v in order to evaluate the **mean speed** v_m given by the following integral:

$$v_m = \int_0^\infty f(v)v\,dv = \sqrt{\frac{8k_B T}{\pi m}} \tag{1.30}$$

At 273.15 K, the mean molecular speed for O_2 is 425 m s^{-1}.

The **root mean square speed** (square root of the mean square speed, *rms*) is given by

$$v_{rms} = \sqrt{\int_0^\infty f(v)v^2\,dv} = \sqrt{\frac{3k_B T}{m}} \tag{1.31}$$

For O_2 at 273.15 K, the *rms* speed is equal to 461 m s^{-1} (**E2**).

1.2.3 Mean Free Path

Consider the mutual diffusion of two gases, oxygen and nitrogen, at a defined pressure and temperature, after the wall separating them has been removed (Fig. 1.7a). When the equilibrium is reached, the gas mixture has the same concentration everywhere in the system. In a container with a few liters of volume, the equilibrium state may be reached in approximately 1 h. This result may look surprising, since molecular speeds are on the order of hundreds of meters per second. While each molecule has a large number of collisions, not all of them drive the molecule away from the starting point, since some collisions drive the molecule backward.

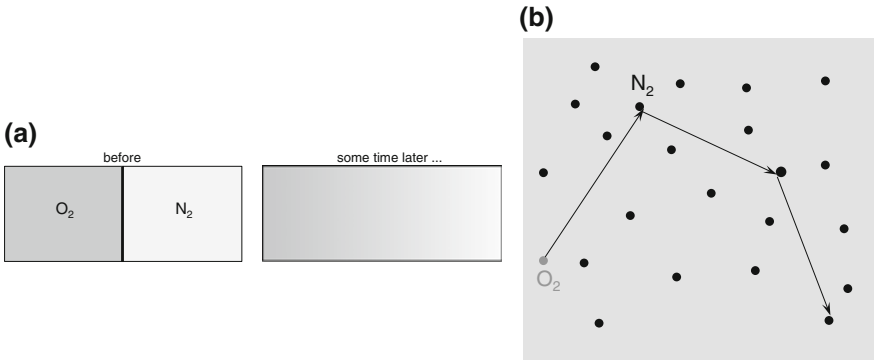


Fig. 1.7 **a** Diffusion of two gases at the same pressure and temperature. **b** Test molecule of O_2 diffusing through N_2. Note that the colors used to distinguish the gases are fictitious: both O_2 and N_2 do not absorb in the visible spectrum

Let us follow an oxygen molecule on its path through nitrogen gas (Fig. 1.7b). To simplify matters, let us assume that the nitrogen molecules are fixed. Through what distance does the oxygen molecule move before it first collides with a nitrogen molecule? How long does it move before its second/third collision? The average distance between collisions is called the **mean free path**.

In this thought experiment, the test molecule O_2 and the molecules N_2 are represented by spheres of the same diameter d. The **cross section for collision** is the area of a circle of radius d (πd^2) (Fig. 1.8a). Between consecutive collisions, this circle spans a cylinder whose height is equal to the distance traveled by the test molecule (Fig. 1.8b). The volume of the broken cylinder is called the interaction volume, and it contains all the nitrogen molecules that collide with the test molecule within a certain time interval Δt.

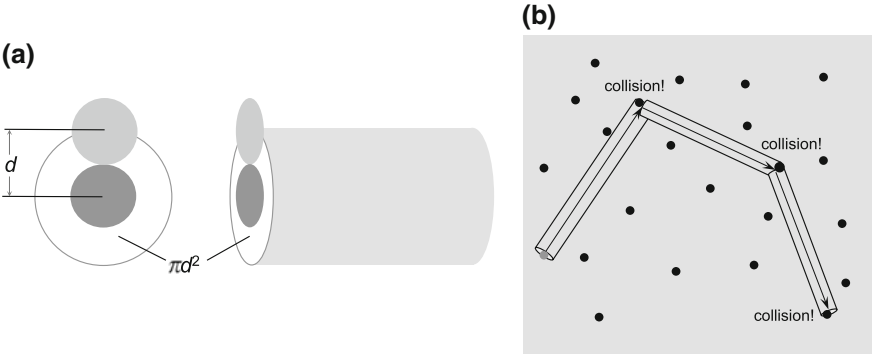


Fig. 1.8 **a** Cross section for collision between spheres of equal radius. **b** All molecules colliding with the test molecule have their centers of mass within the broken cylinder

A reasonable estimate λ for the mean free path can be obtained by dividing the distance δ traveled by the test molecule by the number of collisions N_c, that is, the number of N_2 molecules within the interaction volume (= volume of the broken cylinder),

$$\lambda = \delta / N_c \tag{1.32}$$

N_c is given by

$$N_c = V_{int} \rho_N \tag{1.33}$$

where ρ_N is the number density of molecular nitrogen and the interaction volume V_{int} is

$$V_{int} = \pi d^2 \delta \tag{1.34}$$

Substitution of (1.34) in (1.33) and then in (1.32) leads to

$$\lambda = \frac{\delta}{\pi d^2 \delta \rho_N} \left(\pi d^2 \rho_N\right)^{-1} \tag{1.35}$$

This result is based on the less realistic assumption of the nitrogen molecules remaining fixed while the test molecule travels among them. However, the distance traveled by the test molecule [δ in the numerator of (1.35)] and the length of the interaction cylinder [δ in the denominator of (1.35)] are distinct quantities, which we name δ_n and δ_d, respectively. In fact, the distance traveled by the test molecule δ_n is equal to the *rms* speed times the corresponding time interval,

$$\delta_n = v_{rms} \Delta t \tag{1.36}$$

whereas δ_d, the *real* length of the interaction cylinder, is equal to the *rms* speed of the test molecule *relative* to the speed of the colliding nitrogen molecule,

$$\delta_d = (v_{rel})_{rms} \Delta t \tag{1.37}$$

It can be shown (see **§2**) that

$$(v_{rel})_{rms} = \sqrt{2} v_{rms} \tag{1.38}$$

Hence, substitution of (1.36) in (1.35) and of (1.38) and (1.37) in (1.35) leads to the following expression for the mean free path:

$$\lambda = \frac{\delta_n}{\pi d^2 \delta_d \rho_N} = \frac{v_{rms} \Delta t}{\pi d^2 \sqrt{2} v_{rms} \Delta t \rho_N} = \left(\sqrt{2} \pi d^2 \rho_N\right)^{-1} \tag{1.39}$$

Considering the assumptions of the kinetic model of gases, this expression can give only an order of magnitude for the mean free path. According to (1.39), the mean free path is inversely proportional to the cross section for collisions and the gas number density. In fact, increases of the cross section and of the gas number density are expected to lead to decreases in the mean free path. Note that the mean free path is independent of the distance traveled by the test particle [see (1.39)], since this distance was chosen in an arbitrary way. Considering the approximate nature of (1.39), one can deal with d as a parameter whose value can be chosen so that a realistic value is obtained for the collision cross section. In particular, d = 0.3 nm (= 3 Å) allows one to obtain a reasonable estimate for the mean free path.

The quantity ρ_N can be obtained from the gas pressure and temperature using (1.8), assuming ideal gas behavior. Alternatively, in order to relate ρ_N to the molecular diameter d, one can take the ratio of the liquid and gas densities and equate this ratio to the inverse ratio of the molecular volumes,

$$\frac{\rho_l}{\rho_g} = \frac{m/V_l}{m/V_g} = \frac{V_g}{V_l} \tag{1.40}$$

since a molecule has the same mass m in the liquid and gas phases. For nitrogen, the liquid density is 800 kg m^{-3} and the gas density is 1.2 kg m^{-3}. Therefore, the ratio of densities is approximately equal to 670 ($\approx$ 800/1.2). In the liquid, the volume of one molecule is approximately given by d^3 (volume of a cube containing one nitrogen molecule). Hence, applying (1.40) to nitrogen, we can write

$$670 \approx \frac{V_g}{d^3} \quad \Rightarrow V_g \approx 670 d^3 \tag{1.41}$$

If the definition of number density [see (1.9)] is applied to one molecule in the gas and (1.41) is used, we obtain

$$\rho_N = \left(V_g\right)^{-1} \approx \left(670 d^3\right)^{-1} \tag{1.42}$$

Substitution of (1.42) in (1.39) leads to

$$\lambda = \left(\sqrt{2}\pi d^2 \rho_N\right)^{-1} \approx 670 d^3 / \left(\sqrt{2}\pi d^2\right) \approx 150 d \tag{1.43}$$

Hence, the mean free path is about 150 times the collision cross-sectional radius or the particle diameter (see Fig. 1.8a). The average distance between molecules in the gas is given approximately by

$$\left(670 d^3\right)^{1/3} \approx 10 d \tag{1.44}$$

that is, is about 15 times smaller than the mean free path.

The Maxwell velocity distribution (1.23) characterizes an equilibrium state. When a gas is not at equilibrium, the distribution of velocities differs from (1.23), and the temperature may not be well defined. In a gas, the evolution from an initial velocity distribution to a Maxwell velocity distribution is quite fast, since it takes the time of a few molecular collisions (in most cases, this time is shorter than 10^{-8} s) (**E3**).

1.3 Van der Waals Equation

An ideal gas cannot liquefy by a decrease of temperature or by compression, since the ideal gas model ignores intermolecular interactions. Without these, gases would not liquefy, liquids would not solidify, and liquids and solids could be easily compressed to negligible volumes. In 1873, Van der Waals [1837–1923] derived an empirical, equation based on a modification of the ideal gas equation that accounts for the volume of molecules as impenetrable particles and for attractive interactions that contribute to reducing the pressure of a gas made of particles of nonzero size. He received the Nobel Prize in physics in 1910 for "his work on the equation of state for gases and liquids."

Let us begin by considering the first modification of the ideal gas equation that accounts for the nonzero volume of molecules. For 1 mole of gas,

$$V_{\mathrm{m}} - b \tag{1.45}$$

where V_{m} is the ideal gas molar volume and b accounts for the nonzero volume of molecules per mole of gas, that is, the excluded molar volume. The second modification of the ideal gas equation introduced by Van der Waals accounts for the pressure reduction due to attractive forces between particles. Attraction between particles reduces their velocities, thus decreasing the number of collisions per time unit on a unit area of wall. Considering the $RT/(V_{\mathrm{m}} - b)$ pressure being reduced by δp, we can write

$$p = \frac{RT}{V_{\mathrm{m}} - b} - \delta p \tag{1.46}$$

Van der Waals took δp inversely proportional to the square of the volume,

$$\delta p = \frac{a}{V_{\mathrm{m}}^2} \tag{1.47}$$

Thus, the Van der Waals equation is given by

$$\left(p + \frac{a}{V_{\mathrm{m}}^2}\right)(V_{\mathrm{m}} - b) = RT \tag{1.48}$$

where a and b are **van der Waals constants** (see Table).

Van der Waals constants a and b for selected gases

	a (bar L^2 mol^{-2})	b (L mol^{-1})
Helium	0.0346	0.0238
Neon	0.208	0.0167
Argon	1.355	0.0320
Krypton	2.325	0.0396
Xenon	4.192	0.0516
Carbon dioxide	3.658	0.0429
Carbon disulphide	11.25	0.0726
Ammonia	4.225	0.0371
Water	5.537	0.0305
Methane	2.303	0.0431
Ethane	5.580	0.0651
Hexane	24.84	0.1744
Decane	52.74	0.3043

Values taken from the *Handbook of Chemistry and Physics*, 2011. Alternatively, see Wolfram Alpha for a particular gas

At a specified temperature, the (p, V_m) curve is an **isotherm**. Figure 1.9 shows isotherms for carbon dioxide, from 274 to 344 K, in steps of 10 K. As the temperature of the isotherm increases, the maximum and minimum of the isotherms gradually become closer, until they merge at the **critical temperature**. In the critical isotherm, pressure and molar volume at which maximum and minimum coincide correspond to an inflection called a **critical point**. This point is mathematically defined by the following equalities:

$$\left(\frac{\partial p}{\partial V_m}\right)_T = 0 \quad \left(\frac{\partial^2 p}{\partial V_m^2}\right)_T = 0 \tag{1.49}$$

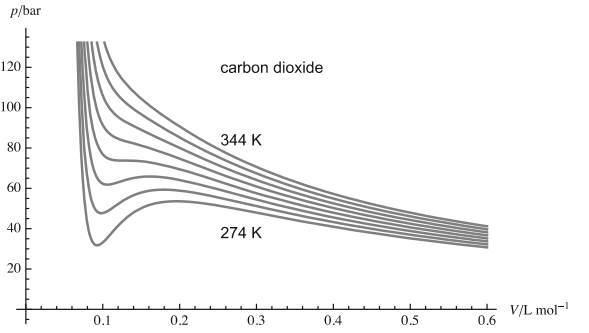


Fig. 1.9 Van der Waals isotherms for carbon dioxide, from 274 to 344 K, in steps of 10 K. Figure obtained with *Mathematica*

Solving this system of equations gives the critical volume V_{mc} and the critical temperature T_c. The corresponding pressure is the critical pressure p_c (see *Mathematica* code **M2**). The critical values depend on the van der Waals constants as follows:

$$V_{mc} = 3b \quad T_c = \frac{8a}{27bR} \quad p_c = \frac{a}{27b^2} \tag{1.50}$$

(**E4**). Note that (1.48) can be written as

$$pV_m^3 - (bp + RT)V_m^2 + aV_m - ab = 0 \tag{1.51}$$

showing that the Van der Waals equation is cubic in V_m. When the linear and quadratic terms cancel, the three roots coincide, meaning that the maximum and minimum of the isotherm converge to the critical point. For $T > T_c$, the isotherms are single-valued, and as the temperature increases, they gradually lose the inflection point and become similar to the ideal gas isotherms (Fig. 1.9) (**E5**).

Figure 1.10 presents a Van der Waals isotherm (continuous solid curve) for carbon dioxide at 270 K. Let us start at E′ in the gas and progress as the volume decreases. The point at which the gas starts to condense is E. When the liquid is in equilibrium with the gas, a progressive volume reduction causes more gas to condense, while the pressure is kept constant during the liquid–gas equilibrium. The dashed straight line AE represents the *real* liquid–gas equilibrium, whereas the oscillating part of the Van der Waals curve that includes a maximum and a minimum is not real, and so the points along the continuous solid curve ABCDE cannot be experimentally observed. Maxwell determined the horizontal part AE and found that the areas α (=ABCA) and β (=CDEC) are equal (**Maxwell construction**). Progressing along the straight line EA as the volume decreases due to gradual condensation of the gas, the point A is attained where all the gas has been condensed to the liquid. Thus, the steep curve AA′ corresponds to the liquid. The steepness of this part of the curve is due to the reduced compressibility of the liquid.

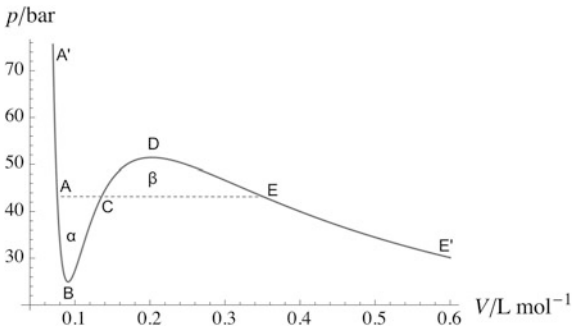


Fig. 1.10 Van der Waals isotherm at 270 K, for carbon dioxide. A′A corresponds to the liquid, the *dashed straight line* ACE corresponds to the liquid–gas equilibrium, and EE′ corresponds to the gas. Graph obtained with *Mathematica*

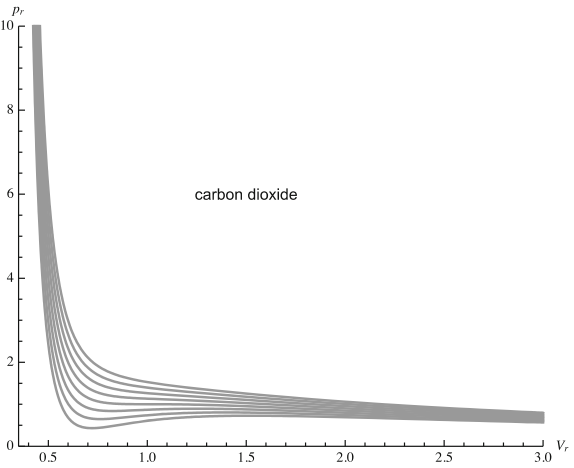


Fig. 1.11 Van der Waals isotherms in reduced variables. The reduced temperatures correspond to the temperatures of the Van der Waals isotherms for carbon dioxide in Fig. 1.9. Figure obtained with *Mathematica*

We now consider the critical values (1.50) and use them as units of volume, temperature, and pressure to obtain the **reduced variables**

$$V_{\mathrm{mr}} = V_{\mathrm{m}}/V_{\mathrm{mc}} \quad T_{\mathrm{r}} = T/T_{\mathrm{c}} \quad p_{\mathrm{r}} = p/p_{\mathrm{c}} \tag{1.52}$$

If the values of V_{m}, T, and p taken from (1.52) are substituted in (1.48), one obtains the Van der Waals equation in reduced variables

$$\left(p_{\mathrm{r}} + \frac{3}{V_{\mathrm{mr}}^2}\right)\left(V_{\mathrm{mr}} - \frac{1}{3}\right) = \frac{8}{3}T_{\mathrm{r}} \tag{1.53}$$

(see *Mathematica* code **M3**). This is a universal equation in the sense that it does not include Van der Waals constants specific to any particular fluid. In fact, comparison of (1.48) with (1.53) shows that a, b, and the gas constant R correspond to 3, 1/3, and 8/3, respectively, that is, these values of a, b, and R lead to unit values of the critical variables. Figure 1.11 shows the Van der Waals isotherms in reduced variables with the reduced temperatures corresponding to the absolute temperatures of the isotherms for carbon dioxide in Fig. 1.9 (see *Mathematica* code **M4**, **E6**).

1.4 Mathematical Tools

In the following sections of the present chapter, most of the expressions use differentials of thermodynamic functions and deal with inexact differentials and line integrals. So the present section has in mind those students who might need some refreshment of mathematical concepts such as exact and inexact differentials, the fundamental theorem of calculus, and line integrals. Those who are aware of these mathematical concepts might skip this section.

1.4.1 Exact Differential

Recall that the **derivative** of a function $f(x)$ at x is given by

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{1.54}$$

The slope of the tangent line at x (Fig. 1.12) is the trigonometric tangent of the angle θ

$$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)} \tag{1.55}$$

where θ is the angle formed by the tangent line and the positive x-axis. Figure 1.12 shows a curve and its tangent at x, and illustrates the difference between

$$\Delta y = y(x + \Delta x) - y(x) \tag{1.56}$$

and

$$\mathrm{d}y = y'(x)\Delta x \tag{1.57}$$

Inspection of Fig. 1.12 also shows that the difference between (1.56) and (1.57), $\Delta y - \mathrm{d}y$, goes to zero as $\Delta x \to 0$. The question is whether $\Delta y - \mathrm{d}y$ approaches zero faster or slower than Δx. To answer this question, we need to take the limit of $(\Delta y - \mathrm{d}y)/\Delta x$ as $\Delta x \to 0$, which is given by

$$\lim_{\Delta x \to 0} \frac{\Delta y - \mathrm{d}y}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} - y'(x) = y'(x) - y'(x) = 0 \tag{1.58}$$

Since $(\Delta y - \mathrm{d}y)/\Delta x$ goes to zero as $\Delta x \to 0$, it can be concluded that $\Delta y - \mathrm{d}y$ approaches zero faster than Δx, and so we can write

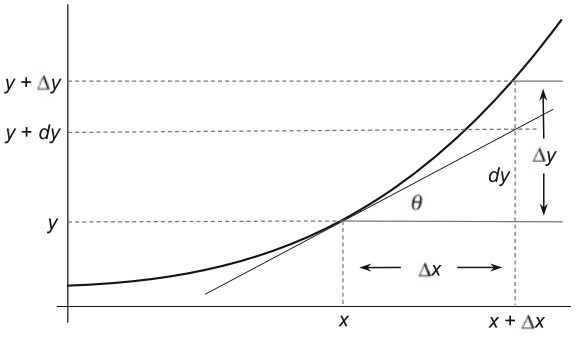


Fig. 1.12 Illustration of Δy and $\mathrm{d}y$

$$\Delta y - \mathrm{d}y = \varepsilon \Delta x \quad \lim_{\Delta x \to 0} \varepsilon = 0 \tag{1.59}$$

where $\varepsilon \Delta x$ goes to zero faster than Δx (for example, as Δx^2). Substitution of (1.57) in (1.59) gives

$$\Delta y = dy + \varepsilon \Delta x = y'(x)\Delta x + \varepsilon \Delta x \tag{1.60}$$

and so

$$\Delta y \approx y'(x)\, \Delta x \tag{1.61}$$

(note the approximation). We now replace Δx and Δy by finite yet arbitrarily small variations $\mathrm{d}x$ and $\mathrm{d}y$ and write

$$\mathrm{d}y = y'(x)\mathrm{d}x \tag{1.62}$$

where $\mathrm{d}y$ is the **differential** of y. Equation (1.62) can be rewritten using the following equivalent forms:

$$\mathrm{d}y = \frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x \quad \mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}x}\mathrm{d}x \tag{1.63}$$

Note that $\mathrm{d}y/\mathrm{d}x$ or the equivalent $\mathrm{d}f/\mathrm{d}x$ represents a limit [see (1.54)], not a quotient of differentials $\mathrm{d}y$ and $\mathrm{d}x$.

For a function $f(x, y)$ of two independent variables, the **partial derivative** of $f(x, y)$ with respect to x is given by

$$f_x = \frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} \tag{1.64}$$

assuming that this limit exists. Note that in evaluating the rate of change of $f(x, y)$ with respect to x, the variable y is held constant. For this reason, we sometimes write $(\partial f/\partial x)_y$ emphasizing that y is kept constant. The partial derivative of $f(x, y)$ with respect to y is given by

$$f_y = \frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} \tag{1.65}$$

Consider now the change Δf as a result of changes in both independent variables, Δx and Δy. We can write the following equalities:

$$\Delta f = f(x+\Delta x, y+\Delta y) - f(x,y)$$
$$\Delta f = [f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y)] + [f(x, y+\Delta y) - f(x,y)]$$
$$\Delta f = \frac{\partial f}{\partial x}\Delta x + \varepsilon_1 \Delta x + \frac{\partial f}{\partial y}\Delta y + \varepsilon_2 \Delta y$$
$$\Delta f = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y$$

If $\varepsilon_1 \to 0$ and $\varepsilon_2 \to 0$ as $\Delta x \to 0$ and $\Delta y \to 0$, then $\Delta f \to \mathrm{d}f$, and we can write

$$\mathrm{d}f = \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y \tag{1.66}$$

where df is the **total differential** of f. For two independent variables x and y, (1.66) is the equivalent of the second equality of (1.63) for one independent variable.

We now consider the differential

$$g(x,y)\mathrm{d}x + h(x,y)\mathrm{d}y \tag{1.67}$$

and ask whether this is an exact differential, that is, whether there is a function $f(x, y)$ such that

$$\mathrm{d}f = \left(\frac{\partial f}{\partial x}\right)_y \mathrm{d}x + \left(\frac{\partial f}{\partial y}\right)_x \mathrm{d}y = g(x,y)\mathrm{d}x + h(x,y)\mathrm{d}y \tag{1.68}$$

that is,

$$\left(\frac{\partial f}{\partial x}\right)_y = g(x,y) \quad \left(\frac{\partial f}{\partial y}\right)_x = h(x,y) \tag{1.69}$$

since dx and dy are arbitrary variations. Note that the first-order derivatives of $f(x, y)$ are related to each other by

$$\frac{\partial\left(\frac{\partial f}{\partial x}\right)}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial\left(\frac{\partial f}{\partial y}\right)}{\partial x} = \frac{\partial^2 f}{\partial x \partial y} \tag{1.70}$$

that is, the **mixed partial second derivatives** are equal. Substitution of (1.69) in (1.70) leads to

$$\frac{\partial g(x,y)}{\partial y} = \frac{\partial h(x,y)}{\partial x} \tag{1.71}$$

which is the requirement for (1.67) to be an **exact differential** (**E7**).

1.4.2 Fundamental Theorem of Calculus

Recall now that the **integral** of a continuous function $f(x)$ in the interval $[a,b]$ is given by

$$\int_b^b f(x)\mathrm{d}x = \lim_{|L|\to 0} \sum_i f(x_i)\Delta x_i \tag{1.72}$$

where $[a,b]$ is divided into n subintervals Δx_i and $|L|$ represents the width of the largest subinterval (Fig. 1.13).

Since $f(x_i)$ times Δx_i is the area of the rectangle of height $f(x_i)$ and width Δx_i, the summation on the right-hand side of (1.72) corresponds to the area of all the rectangles, and its limit is the area under the curve in the interval $[a,b]$ (Fig. 1.13). If the upper limit b is replaced by the variable x defined within the interval $[a,b]$, one obtains the **indefinite integral**

$$F(x) = \int_a^x f(u)du \tag{1.73}$$

where the variable of integration is denoted by u to prevent confusion with the integral's upper limit variable x. It can be shown that

$$F'(x) = f(x) \tag{1.74}$$

i.e., $F(x)$ is the **antiderivative** of $f(x)$. In order to understand this important result, known as the **fundamental theorem of calculus**, we apply the definition of the derivative of $F(x)$ and use (1.73), obtaining

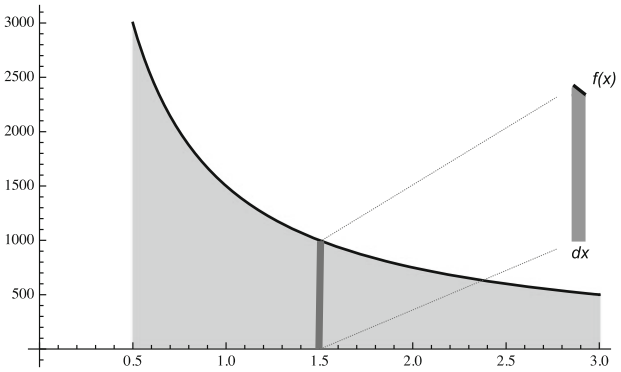


Fig. 1.13 The integral of the function in the interval [0.5,3.0] is equal to the area under the curve. Figure obtained with *Mathematica*

$$F(x+\Delta x)-F(x)=\int_a^{x+\Delta x} f(u)\mathrm{d}u-\int_a^{x} f(u)\mathrm{d}u=\int_x^{x+\Delta x} f(u)\mathrm{d}u=f(\xi)\Delta x$$

$$F(x+\Delta x)-F(x)=f(\xi)\Delta x$$

$$\lim_{\Delta x\to 0}\frac{F(x+\Delta x)-F(x)}{\Delta x}=f(x)\quad \text{i.e.}\quad F'(x)=f(x)$$

The last equality of the first line above shows that $f(\xi)$ is the mean value of $f(x)$ in the interval Δx. When applied to a general interval $[a,b]$, this equality expresses the so-called mean value theorem for integrals. In turn, comparison of the second and third equalities above allows one to conclude that $f(\xi) \to f(x)$ as $\Delta x \to 0$.

1.4.3 Line Integral

Consider now a curve C parametrically described by the position vector of a particle:

$$\boldsymbol{r}(t)=x(t)\,\boldsymbol{i}+y(t)\,\boldsymbol{j}+z(t)\,\boldsymbol{k} \tag{1.75}$$

Without loss of generality, we can assume that the parameter t is the time variable. The particle is driven by a force $\boldsymbol{F}$ over the path defined by the position vector $\boldsymbol{r}(t)$. When the particle undergoes the displacement $\mathrm{d}\boldsymbol{r}$ during the time interval $\mathrm{d}t$, the infinitesimal work $\mathrm{d}W$ done by $\boldsymbol{F}$ on the particle is given by the scalar product $\boldsymbol{F}.\mathrm{d}\boldsymbol{r}$ and the integral of $\mathrm{d}W$ over the curve C is

$$W=\int \mathrm{d}W=\int_C \boldsymbol{F}.\mathrm{d}\boldsymbol{r}=\int_{t_\mathrm{i}}^{t_\mathrm{f}} \boldsymbol{F}.\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\mathrm{d}t \tag{1.76}$$

where t_i and t_f are the limits of integration (the initial and final values of t), and $\boldsymbol{r}(t_\mathrm{i})$ and $\boldsymbol{r}(t_\mathrm{f})$ denote the initial and final positions of the particle on the curve C. The last equality of (1.76) results from

$$\mathrm{d}\boldsymbol{r}=\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\mathrm{d}t \tag{1.77}$$

[see (1.63)]. From the last member of (1.76), the curve can be seen as consisting of a finite number of infinitesimal segments, each of which has a tangent given by $\mathrm{d}\boldsymbol{r}/\mathrm{d}t$ defined at each point of the curve. The direction of the tangent varies continuously with t. Since the integral (1.76) depends on the initial and final points of the path and on the path itself, it is called a **line integral** or **path integral**. The path-dependence of (1.76) is reflected by the scalar product $\boldsymbol{F}\,.\,\mathrm{d}\boldsymbol{r}/\mathrm{d}t$, since both $\boldsymbol{F}$ and $\mathrm{d}\boldsymbol{r}/\mathrm{d}t$ are path-dependent functions. Here $\mathrm{d}W$ is an **inexact differential**.

Exact differentials have an important property: the integral of an exact differential depends only on the initial and final points of integration (the limits of integration),

$$\Delta f = \int_A^B \mathrm{d}f = f(B) - f(A) \tag{1.78}$$

If the integration of an exact differential is done over a cycle (Fig. 1.14), then the initial and final points of integration coincide, and

$$\oint \mathrm{d}f = 0 \tag{1.79}$$

From Fig. 1.14 and (1.79), we can write

$$\oint \mathrm{d}f = \int_{A \to B(1)} \mathrm{d}f + \int_{B \to A(2)} \mathrm{d}f = 0 \tag{1.80}$$

Where

$$\int_{A \to B(1)} \mathrm{d}f = -\int_{B \to A(2)} \mathrm{d}f = \int_{A \to B(2)} \mathrm{d}f \tag{1.81}$$

The equality between the first and the last members of (1.81) allows one to conclude that the integral of an exact differential is path-independent, because the paths 1 and 2 that link A and B were *arbitrarily* chosen. Then it can be concluded that (1.79) is a necessary and sufficient condition for df to be an exact differential [*necessary condition*: if df is an exact differential, then (1.79) is valid; *sufficient condition*: if (1.79) is valid, then df is an exact differential].

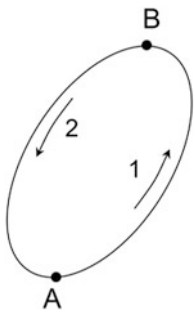


Fig. 1.14 The integral of an exact differential over a cycle is zero

1.5 Thermodynamic Systems

Every thermodynamic study begins with a clear distinction between the object of study, the **system**, and everything else, the **exterior**. Thus, the system is absolutely central to every thermodynamic study. One can say that the system is the "protagonist" in thermodynamics. It may be a gas, two metal blocks in contact, a liquid and its gas in equilibrium, a solute in a solution, a chemical reaction. Besides the system composition, a thermodynamic study requires the knowledge of its boundaries, since interactions may occur between the system and its surroundings. On the whole, system and exterior form the **universe of the transformation**.

Systems can be classified into three types, *isolated*, *closed*, and *open*, according to the way they interact with the exterior. An **isolated system** does not exchange matter or energy with the exterior (Fig. 1.15). The exterior of an isolated system is irrelevant, and it may be considered nonexistent. Thus, an isolated system coincides with the universe of the transformation. For this reason, isolated systems have considerable theoretical importance, in particular in the formulation of the first and second laws of thermodynamics, since these laws become simpler and more comprehensive when they are referred to isolated systems. However, the complete and effective isolation of a system is difficult to realize, since it requires walls that behave as perfect insulators. A **closed system** does not exchange matter with its surroundings, and an **open system** exchanges matter and energy with its surroundings (Fig. 1.15).

When the variables needed to characterize a system take *defined* values, the system is said to exist in a thermodynamic **state**, and the variables are called **state variables**. Pressure p, temperature T, volume V, and amount of substance n_k (the index k distinguishes different substances) are state variables used in describing equilibrium states. Variables such as volume and amount of substance provide information about the size of a system and for that reason are called **extensive variables**. On the other hand, variables such as pressure and temperature are local variables. As such, they do not provide information on the size of the system and are called **intensive variables**. The ratio of two extensive variables that have the same size-dependence is an intensive variable. For example, the ratio of mass to volume is density, an intensive variable.

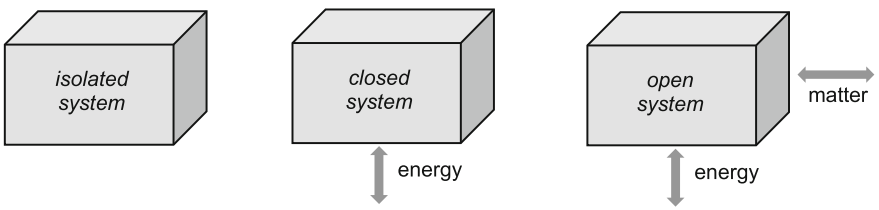


Fig. 1.15 Systems can be classified according to the way in which they interact with the exterior

An **equilibrium state** is a time-independent state. An isolated system away from equilibrium inevitably evolves to an equilibrium state through **irreversible processes**. Examples of these are heat conduction, matter diffusion, and chemical reactions. When a system is not at equilibrium, additional information is required on the time-dependence of the irreversible processes that occur in the system before it reaches the equilibrium state. During an irreversible process, the values of the state variables are time-dependent, and time is an external variable (a parameter) measured by a chronometer. In a chemical reaction, the amounts of reactants and products are related by the reaction stoichiometry, and it is possible to define a single state variable, called the **extent of the reaction**, that accounts for the time evolution of the chemical reaction toward equilibrium.

The first and second laws of thermodynamics involve the energy U and entropy S, and these thermodynamic **state functions** can be expressed as functions of the state variables temperature, volume, and amount of substance:

$$U = U(V, T, n_k) \quad S = S(V, T, n_k) \tag{1.82}$$

As we will find later on in this chapter, scientific convenience leads to the definition of new state functions using the energy U, entropy S, and state variables p, T, V, and n_k. Each thermodynamic state is characterized by the corresponding set of values of the state variables.

1.6 Heat and Work

The intensive variables temperature and pressure are associated with two fundamental types of equilibrium, namely thermal equilibrium and mechanical equilibrium. Consider an isolated system formed by two blocks of the same metal, at temperatures T_1 and T_2, with $T_1 > T_2$ ($= T_1 - \mathrm{d}T$) (Fig. 1.16). Heat $\mathrm{d}Q$ flows spontaneously from the part of the system at temperature T_1 to the part of the system at temperature T_2, until both parts reach the same temperature. One says that the system has then attained **thermal equilibrium**. It is assumed that small variations in volume (part 1 reduces its volume as it cools; part 2 increases its volume as it warms) can be easily accommodated by the system, since the heat transferred is

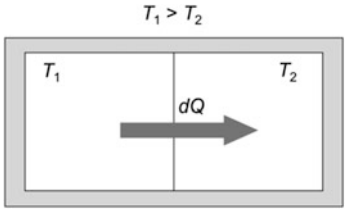


Fig. 1.16 Heat flows from higher to lower temperature

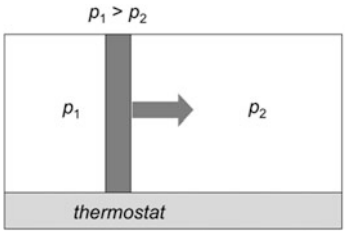


Fig. 1.17 At constant temperature, a frictionless wall moves from higher to lower pressures

small. Note that a system in thermal equilibrium has the same temperature in all points.

We now consider a closed system at constant temperature formed by a gas in two compartments with pressures p_1 and p_2, and $p_1 > p_2 (= p_1 - \mathrm{d}p)$. A frictionless wall separates these compartments and moves spontaneously from higher to lower pressures (Fig. 1.17). When the pressures become equal, we say that the system has attained **mechanical equilibrium**. A system in mechanical equilibrium has the same pressure at all points.

1.6.1 Mechanical Work

Consider now a gas at constant temperature in a cylindrical container with a frictionless wall (Fig. 1.18a). Assume that the initial gas pressure is greater than the external pressure, $p > p_{\mathrm{ext}}$. Then the gas volume increases, and its pressure decreases until the pressure becomes the same on both sides of the frictionless wall. The gas undergoes an isothermal expansion (an increase in volume) and performs mechanical work on the surroundings. The work is assigned the sign of the energy

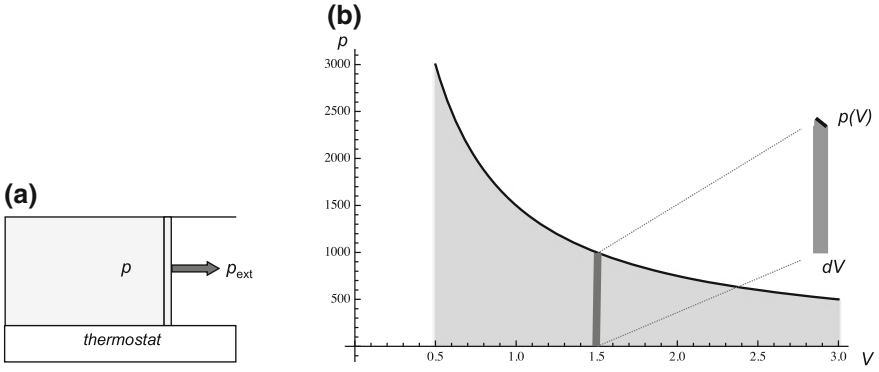


Fig. 1.18 **a** The system does work on the exterior during an isothermal expansion. **b** Apart from its sign, the mechanical work is equal to the area under the curve $p(V)$

change produced on the system, that is, the work done on expansion is negative. In turn, the work done on compression is positive, just like the corresponding energy change of the system (on compression, the energy of the system increases). For an infinitesimal quantity of work, one has

$$\mathrm{d}W = -F\mathrm{d}x = -\frac{F}{A}A\mathrm{d}x = -p\mathrm{d}V \tag{1.83}$$

where F is the force exerted on the moving wall or piston and A is the area of the piston. Integration of (1.83) for the whole expansion leads to

$$W = -\int_{V_\mathrm{i}}^{V_\mathrm{f}} p\mathrm{d}V \tag{1.84}$$

where V_i and V_f are the initial and final volumes, respectively. The absolute value of the work is the area under the curve $p(V)$ (Fig. 1.18b). Work is a **path-dependent** quantity, or **process-dependent** quantity, not a state function (**E8**). Mathematically, $\mathrm{d}W$ is an inexact differential.

1.7 First Law

Can energy be obtained from "nothing" or simply destroyed to give "nothing"? In the past, there were many attempts to obtain mechanical work for free. The idea was based on so-called perpetual motion machines (Angrist 1968). However, none of those machines were really successful. Indeed, all the experimental knowledge gathered so far has confirmed that energy changes should always be accounted for. Even in the field of nuclear reactions, where energy and mass are interchangeable ($E = mc^2$), energy conservation has been always confirmed overall.

If one considers the universe (= system + exterior) of a transformation, the total energy is conserved, that is, the energy change for any infinitesimal transformation is zero:

$$\mathrm{d}U_\mathrm{univ} = 0 \tag{1.85}$$

For an isolated system, the exterior is irrelevant, because an isolated system does not exchange matter or energy with its surroundings. Therefore,

$$\mathrm{d}U_\mathrm{isol} = 0 \tag{1.86}$$

In a closed system, the system can exchange energy with the surroundings in two ways, through heat flow and through work. Representing by $\mathrm{d}Q$ the heat exchanged between the system and its surroundings during the time interval $\mathrm{d}t$ and by $\mathrm{d}W$ the

work exchanged between the system and its surroundings during the same time interval, then

$$\mathrm{d}U = \mathrm{d}Q + \mathrm{d}W \tag{1.87}$$

Since dW is an inexact differential, dQ has to be an inexact differential, because only the sum of two inexact differentials can give the exact differential dU. Thus, every energy change in U depends only on the initial and final states of the transformation, not on the path followed to convert the initial state into the final state, and so

$$\oint \mathrm{d}U = 0 \tag{1.88}$$

The dQ and dW terms in (1.87) are positive if they correspond to heat and work transferred *from the exterior to the system* (in this case, the system gains energy), and negative if they correspond to heat or work transferred *from the system to the exterior* (in this case, the system loses energy). This convention of signs for dQ and dW is consistent with the system being the "protagonist" in thermodynamics.

Substitution of (1.83) in (1.87) leads to

$$\mathrm{d}U = \mathrm{d}Q - p\mathrm{d}V \tag{1.89}$$

At constant volume, the mechanical work is zero, and

$$\mathrm{d}Q_\mathrm{V} = \mathrm{d}U \tag{1.90}$$

where dQ_V represents the exchanged heat at constant volume (Fig. 1.19). At constant pressure,

$$\mathrm{d}Q_\mathrm{p} = (\mathrm{d}U + p\mathrm{d}V)_\mathrm{p} = (\mathrm{d}U + p\mathrm{d}V + V\mathrm{d}p)_\mathrm{p} = [\mathrm{d}(U + pV)]_\mathrm{p} \tag{1.91}$$

where the added term $(V\mathrm{d}p)_\mathrm{p}$ does not affect the result, because at constant pressure, dp = 0, and so Vdp = 0. If we define the state variable **enthalpy** (*en-* = within; *thalpein* = to heat),

$$H \equiv U + pV \tag{1.92}$$

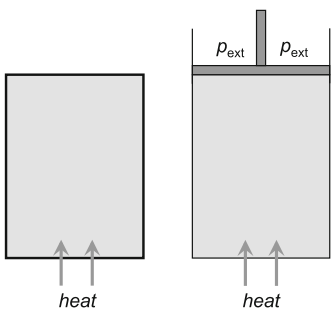


Fig. 1.19 Heat at constant volume (*left*) and heat at constant pressure (*right*)

then substitution of (1.92) in (1.91) leads to

$$\mathrm{d}Q_\mathrm{p} = \mathrm{d}H \tag{1.93}$$

Note that U is a state function, both p and V are state variables, and the product and sum of two state variables is a state function. This reasoning allows to conclude that $U + pV$ is a state function.

1.7.1 Heat Capacities

Expressions (1.90) and (1.93) can be used to define **isochoric heat capacity** (heat capacity at constant volume, *isochoric* from Greek *isos "equal"* + *choros "space"*),

$$C_\mathrm{V} = \frac{\mathrm{d}Q_\mathrm{V}}{\mathrm{d}T} = \left(\frac{\partial U}{\partial T}\right)_V \tag{1.94}$$

and **isobaric heat capacity** (heat capacity at constant pressure, *isobaric* from Greek *isos "equal"* + *baros "weight"*),

$$C_\mathrm{p} = \frac{\mathrm{d}Q_\mathrm{p}}{\mathrm{d}T} = \left(\frac{\partial H}{\partial T}\right)_p \tag{1.95}$$

(**E9**, **E10**). Since heat is an extensive variable and temperature is an intensive variable, heat capacity is an extensive variable (the ratio of one extensive variable to one intensive variable). Tabulated values found in books of data are values of intensive quantities (extensive quantities are *sample*-dependent quantities). For that reason, it is useful to define

$$c = C/m \tag{1.96}$$

where c is the **specific heat capacity** and m is the mass, usually expressed in grams. Thus, the specific heat capacity is the heat capacity for 1 g of substance. Whenever it is possible and appropriate to deal with an amount of substance, then

$$C_\mathrm{m} = C/n \tag{1.97}$$

where C_m is the **molar heat capacity** and n is the amount of substance. Figure 1.20 shows the molar heat capacities at constant pressure (1 bar = 100,000 Pa), for two important gases in the atmosphere, water and carbon dioxide, as functions of temperature. At ambient temperature and pressure of 1 atm, the heat capacity for liquid water approximately doubles the heat capacity for gaseous water.

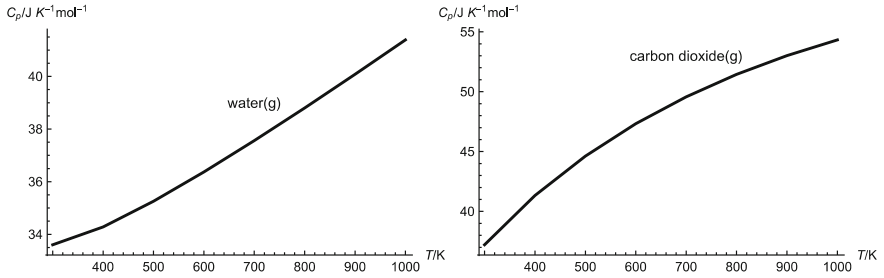


Fig. 1.20 Molar heat capacities at constant pressure, for water and carbon dioxide in the gaseous states, as functions of temperature. Values taken from the *Handbook of Chemistry and Physics* 2011, and graphs obtained with *Mathematica*

1.7.2 Calorimeter

The heat released or absorbed in a transformation can be experimentally determined with a **calorimeter**. Usually, the transformation is made to occur inside a chamber, at constant volume, and the heat associated with the transformation is determined from the change in the temperature of the surrounding water. Derived from (1.94), the basic equation is

$$\mathrm{d}Q_\mathrm{V} = C_\mathrm{cal}\mathrm{d}T \tag{1.98}$$

where C_cal is the **heat capacity of the calorimeter**. In order to determine $\mathrm{d}Q_\mathrm{V}$ from a measured $\mathrm{d}T$, C_cal has to be known beforehand. To this end, one chooses to

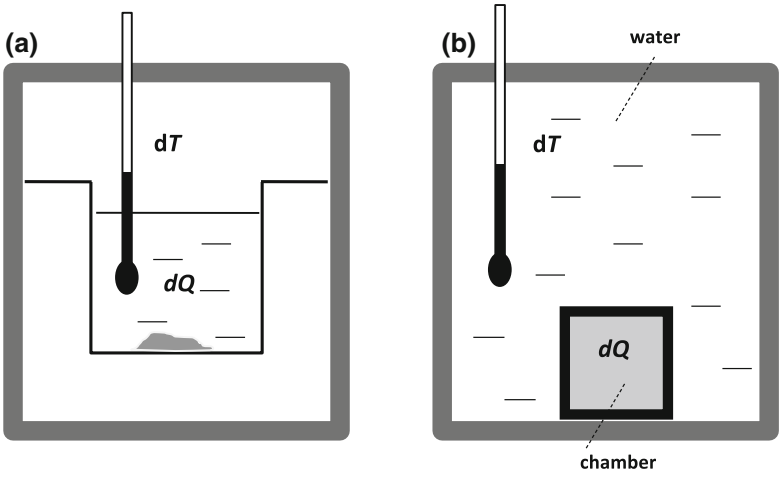


Fig. 1.21 Schematic representation of calorimeters for determining the heat of a saturated solution (*at left*) and the heat of combustion of a compound (*at right*)

measure dT for a known dQ_V, which can be produced by a current-carrying resistor. For a current intensity I and a resistance R, the voltage across the resistor is $V = IR$ and the heat evolved by the resistance in the time interval dt is I^2Rdt. Once C_{cal} is known (note that C_{cal} is a calorimeter constant), dQ_V can be determined from the change in temperature dT measured by the same calorimeter.

A calorimeter should be suitable to the type of transformation whose heat is to be determined (Fig. 1.21). For example, the heat of combustion of a compound is determined in a **bomb calorimeter**, with the combustion being carried out inside a chamber pressurized to about 20 atm with pure oxygen to ensure that the combustion is complete.

When the heat to be determined is smaller than 1 J (for example, in the thermodynamic study of biological systems), the sample that absorbs or releases heat is kept at constant temperature by a very sensitive heat exchanger that measures the exchanged heat. This type of calorimeter is called an **isothermal calorimeter**.

1.7.3 Standard States

The thermodynamic properties of a system depend on the nature and physical state (gas, liquid or pure crystalline solid) of its component substances, as well as on the physical variables such as temperature, pressure, and, in the case of solutions, concentrations of solutes (partial pressures for gaseous mixtures). Thus, tabulated data on a system, besides specifying its temperature, rely on a standard pressure p^0 and, in the case of solutions, on a standard molality m^0 or a standard molarity c^0 for each solute. Without such standard values, tabulated thermodynamic data could be of little practical use, due to the dispersion of physical conditions. The most common choice of tabulating data includes a reference temperature equal to 298.15 K, the pressure $p^0 = 1$ bar (= 100,000 Pa), and, in the case of solutions, the solute molality $m^0 = 1$ mol of solute per kg of solvent, or the solute molarity $c^0 = 1$ mol of solute per dm^3 of solution. The **standard state** of a pure substance at a specified temperature is its most stable state (gas, liquid, or pure crystalline solid) at the standard pressure $p^0 = 1$ bar.

In the **gaseous state**, the standard state is the *hypothetical state* in which the gas has ideal behavior at the standard pressure p^0. This means that a real gas at 1 bar of pressure is not in its standard state. Deviations from ideal behavior result from intermolecular interactions that depend on the gas, its pressure, and temperature, and so would not lead to a definable standard state.

The standard state of a **pure liquid** or a **pure solid** is its state at the standard pressure p^0. Unlike gases, liquids and solids have reduced compressibility, since their volumes are little affected by pressure variation.

The standard state of a **solute** is its *hypothetical state* in a solution with ideal behavior, at the standard concentration m^0 or c^0, and the standard pressure p^0. Ideal behavior corresponds to the absence of solute–solute intermolecular interactions, implying that the solute vapor pressure is proportional to the mole fraction of the solute in solution (Henry's law). At standard concentration m^0 or c^0 and standard

pressure p^0, the solute in a real solution is not in the standard state, since the deviations from ideal behavior are solution-specific, that is, do not lead to a definable standard state.

Absolute energies cannot be determined. Only energy changes can be measured or evaluated. This applies to the energy U, the enthalpy H, and other thermodynamic state functions later considered in this chapter.

The enthalpies of formation of *all elemental substances* in their standard states are zero, at *all temperatures*. Hence, at a specified temperature, the **standard molar enthalpy of formation** H_f^o of a compound is the difference between its molar enthalpy in its standard state and the enthalpies of its constituent elements in their standard states. In addition, at a specified temperature, the **standard reaction enthalpy** H_r^o is the difference between the stoichiometric sum of the enthalpies of formation of the products and the stoichiometric sum of the enthalpies of formation of the reactants, with reactants and products being in their standard states (**E11**). Tabulated data of standard thermodynamic properties at $T = 298.15$ K usually include data for a large number of compounds (see *Handbook of Chemistry and Physics*, 2011).

1.8 Reversible Heat Engine

Can heat be converted into work with 100 % efficiency? All existing experimental knowledge gives a negative answer to this question. Even in idealized conditions of maximum efficiency, a fraction of the heat is necessarily released to a heat sink without being converted into work. In contrast, work can be converted into heat with 100 % theoretical efficiency. In order to understand the consequences of this asymmetry of nature, we consider the **heat engine** (a device that converts heat into work) functioning under idealized conditions of maximum efficiency.

The first thermodynamic studies on the heat engine were strongly impelled by the great impact of the steam engine on charcoal extraction, transportation, agriculture, and industry during the so-called industrial revolution, from the second half of the eighteenth century to the first half of the nineteenth. In particular, the studies carried out by Sadi Carnot (1796–1832) led to an understanding of the conditions for maximum efficiency of the heat engine.

Under idealized conditions of maximum efficiency, the essential workings of the heat engine are schematically shown in Fig. 1.22. During a complete cycle, the heat engine absorbs heat Q_h from the **heat source** (hot reservoir) at temperature T_h, converts part of Q_h into work W, and discards the remaining heat Q_c ($= Q_h - W$) to the **heat sink** (cold reservoir) at temperature T_c (note that Q_h, W, and Q_c represent absolute values). After performing these actions, the heat engine returns to its initial state to complete a cycle. Hence, the variations of all state variables in the heat engine (our system) are zero. The temperatures of the heat source and heat sink are not affected by the release or capture of the heats Q_h and Q_c, respectively. Hence,

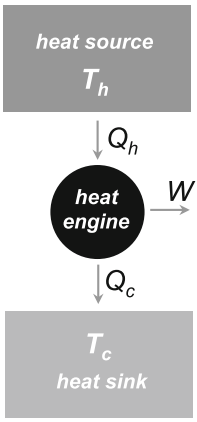


Fig. 1.22 Essential workings of an idealized heat engine

the heat source and the heat sink are assumed to be infinite heat reservoirs, that is to say, to have infinite heat capacities. The **efficiency** of the heat engine is given by

$$\eta = \frac{W}{Q_\mathrm{h}} \tag{1.99}$$

According to the first law,

$$Q_\mathrm{h} - Q_\mathrm{c} = W \tag{1.100}$$

Substitution of (1.100) in (1.99) leads to

$$\eta = 1 - \frac{Q_\mathrm{c}}{Q_\mathrm{h}} \tag{1.101}$$

Thus, the heat engine efficiency is a positive quantity smaller than 1. Maximum efficiency corresponds to the minimum value of $Q_\mathrm{c}/Q_\mathrm{h}$, that is, to the minimum value of Q_c for the same value Q_h. The maximum efficiency of the heat engine cannot depend on the engine materials or engine gears, since otherwise, the efficiency could in principle be surpassed by an improved choice of these items. This reasoning leads to the conclusion that the maximum efficiency of a heat engine can depend only on the temperatures of the heat source and heat sink, τ_h and τ_c,

$$\eta_\mathrm{max} = 1 - \frac{Q_\mathrm{c}}{Q_\mathrm{h}} = 1 - f(\tau_\mathrm{c}, \tau_\mathrm{h}) \tag{1.102}$$

Carnot's theorem. Note that under the conditions of maximum efficiency, volume changes associated with mechanical work cannot produce thermal gradients, for otherwise, additional heat would be released to the heat sink, and maximum efficiency would not be attained. Hence, the heat engine with maximum efficiency takes unlimited time to accomplish its operations. Only then can the heat engine

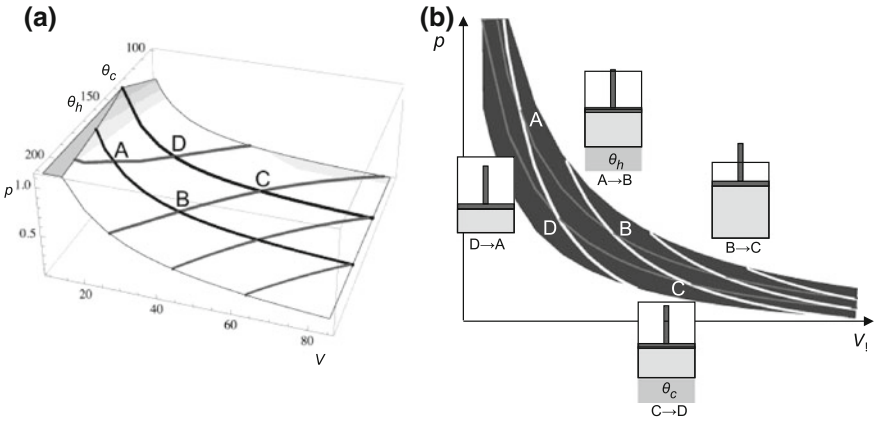


Fig. 1.23 A Carnot engine cycle is formed by an isothermal expansion (*AB*), an adiabatic expansion (*BC*), an isothermal compression (*CD*), and an adiabatic compression (*DA*). **a** $p = R\theta/V$ surface with isothermal and adiabatic curves. **b** Surface (**a**) projected in the plane (p, V, 100) seen from the point (p, V, θ) = (0, 10, 200). Ideal gas surface with isotherms and adiabatic curves obtained with *Mathematica*

follow a reversible path and attain both thermal and mechanical equilibrium at each point in the path, thus being a reversible heat engine.

1.8.1 Carnot's Heat Engine

Carnot's heat engine consists of an ideal gas that operates over a reversible cycle formed by steps $A \rightarrow B$ (isothermal expansion), $B \rightarrow C$ (adiabatic expansion; see **§3**), $C \rightarrow D$ (isothermal compression), and $D \rightarrow A$ (adiabatic compression) (Fig. 1.23) (see *Mathematica* code **M5**). The ideal gas temperatures for the heat source and the heat sink are represented by θ_h and θ_c, respectively.

It can be shown (see **§4**) that the Carnot engine's efficiency depends only on the temperatures of the heat source and heat sink according to the expression

$$\eta_\mathrm{Carnot} = 1 - \frac{Q_\mathrm{c}}{Q_\mathrm{h}} = 1 - \frac{\theta_\mathrm{c}}{\theta_\mathrm{h}} \tag{1.103}$$

where θ_h and θ_c are the temperatures given by the ideal gas empirical scale of temperatures, (**E12**).

1.8.2 Absolute Temperature

As shown before, the heat engine's efficiency is a positive quantity less than 1 [see (1.101)]. For a reversible heat engine, the efficiency depends only on the

temperatures of the hot and cold reservoirs. These two conclusions are combined in (1.102), which, in turn, leads to

$$\frac{Q_c}{Q_h} = f(\tau_c, \tau_h) \tag{1.104}$$

This equality applies to a reversible heat engine operating between temperatures τ_h and τ_c (Fig. 1.24a). Consider now two reversible heat engines, one operating between temperatures τ_h and τ', and the other operating between temperatures τ' and τ_c (Fig. 1.24b). The reservoir at the intermediate temperature τ' operates as a cold reservoir for the first reversible heat engine and absorbs the amount of heat Q', and operates as a hot reservoir for the second reversible heat engine, releasing the same amount of heat Q'. Adding W' (= $Q_h - Q'$) to W'' (= $Q' - Q_c$) cancels Q' and gives the work W (= $W' + W'' = Q_h - Q_c$). From (1.104), we can write

$$\frac{Q_c}{Q'}\frac{Q'}{Q_h} = f(\tau_c, \tau')f(\tau', \tau_h) = \frac{f(\tau_c, \tau')}{f(\tau_h, \tau')} = \frac{g(\tau_c)}{g\tau_h} \tag{1.105}$$

where the second equality results from $Q'/Q_h = 1/(Q_h/Q') = 1/f(\tau_h, \tau')$.

The reversible heat engines of Fig. 1.24 receive the same amount of heat Q_h from the heat source at τ_h, release the same amount of heat Q_c to the heat sink at τ_c, produce the same amount of work $W = W' + W''$ thus having the same efficiency. Therefore, the intermediate temperature τ' is irrelevant and can be omitted as an argument of the function f, thus justifying the last equality of (1.105). In addition, we can assume that the function $g(\tau)$ takes the simple form of a constant c times τ and define a temperature scale for which $c = 1$, i.e.,

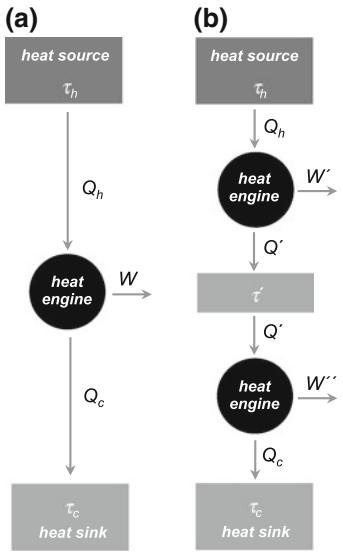


Fig. 1.24 Reversible heat engines: the efficiencies of (**a**) and (**b**) are equal

$$g(\tau) = c\,\tau \equiv T \tag{1.106}$$

From (1.104) and (1.105), it can be concluded that the functions f and g can have only positive values. Thus, the identity of (1.106) defines a **scale of absolute temperatures**. Substitution of (1.106) in (1.105) and then in (1.102) leads to

$$\eta_{\max} = 1 - \frac{Q_c}{Q_h} = 1 - \frac{T_c}{T_h} \tag{1.107}$$

(see *Mathematica* code **M6**). Rearrangement of (1.107) leads to the following equality for a reversible cycle:

$$\frac{Q_h}{T_h} + \frac{-Q_c}{T_c} = 0 \tag{1.108}$$

Comparison of (1.107) with (1.103) allows on to conclude that the scales of the ideal gas temperature and absolute temperature coincide.

Consider now a heat engine that operates in a cycle between temperatures T_h and T_c and absorbs Q_h from the heat source but produces less work than the corresponding Carnot engine ($W^{irr} < W$), since volume changes during production of mechanical work are associated with thermal gradients. These volume changes lead to irreversible heat discarded to the heat sink ($Q_c^{irr} > Q_c$). For this irreversible cycle,

$$\frac{Q_h}{T_h} + \frac{-Q_c^{irr}}{T_c} < 0 \tag{1.109}$$

1.9 Entropy and the Second Law

The equality for a Carnot cycle (1.108) can be generalized to an arbitrary reversible cycle. To this end, consider a reversible cycle and the enclosed area occupied by a mesh of Carnot cycles in which the temperatures of the isotherms differ by ΔT. This mesh comprises two types of Carnot cycle (Fig. 1.25): those that are completely inside the arbitrary cycle and share their isothermal and adiabatic transformations with four neighboring Carnot cycles, and those that are over the cyclic path and have fewer than four neighboring Carnot cycles. Transformations that are shared by two contiguous Carnot cycles are carried out in opposite directions in both cycles and cancel out. The isothermal and adiabatic transformations that are not canceled describe a zigzag path over the periphery of the arbitrary cycle (Fig. 1.25). In the limit of infinitesimal Carnot cycles (the difference between temperatures of consecutive isotherms tends to zero, $\Delta T \to 0$, and consecutive adiabatics gradually get closer), the zigzag path becomes indistinguishable from an **arbitrary reversible cycle**. Hence, the generalization of (1.108) leads to

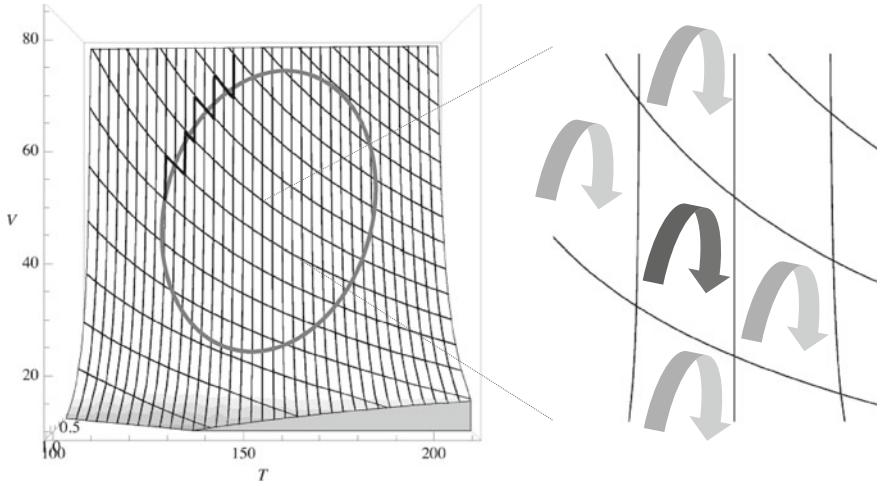


Fig. 1.25 An arbitrary reversible cycle can be covered to the required precision by a mesh of infinitesimal Carnot cycles. Carnot cycles are carried out in the same direction (in this figure, they are performed in the clockwise direction). Figure obtained with *Mathematica*

$$\frac{Q_h}{T_h} + \frac{-Q_c}{T_c} = 0 \quad \Rightarrow \oint \frac{dQ_{rev}}{T} = 0 \tag{1.110}$$

Unlike Q_h and Q_c, which represent absolute values, dQ_{rev} includes a positive or negative sign for absorbed or released heat, respectively. The last equality of (1.110) shows that dQ_{rev}/T is an exact differential [see (1.79)], that is, dQ_{rev}/T is the differential of a state function, which was named by Rudolf Clausius (1822–1888) **entropy** (*en "inside"* + from Greek *tropé "transformation"*) and is usually represented by a capital S:

$$\oint \frac{dQ_{rev}}{T} = 0 \quad \oint dS = 0 \quad dS = \frac{dQ_{rev}}{T} \tag{1.111}$$

(**E13**, **E14**). In turn, generalization of (1.109) for an **arbitrary irreversible cycle** leads to

$$\frac{Q_h}{T_h} + \frac{-Q_c^{irr}}{T_c} < 0 \quad \Rightarrow \oint \frac{dQ}{T} < 0 \tag{1.112}$$

Formulas (1.111) and (1.112) can be combined as follows:

$$\oint \frac{dQ}{T} \leq 0 \quad \oint dS = 0 \quad \therefore dS \geq \frac{dQ}{T} \tag{1.113}$$

where the inequalities hold for irreversible processes. An irreversible process occurs in a specific direction, not in its reverse, and evolves to equilibrium. Examples of irreversible processes are the flow of heat from a higher to a lower temperature, the motion of a piston from a higher to a lower pressure, the melting of an ice cube at ambient temperature, the fall of a weight from a higher to a lower altitude.

The last inequality of (1.113), called **Clausius's inequality**, shows that dS is greater than dQ/T when the latter corresponds to an irreversible exchange of heat between system and surroundings. Hence, for an irreversible process, the difference between dS and dQ/T is *always* greater than zero,

$$\mathrm{d}S - \frac{\mathrm{d}Q}{T} > 0 \tag{1.114}$$

According to Ilya Prigogine (1917–2003, Nobel Prize in chemistry in 1977), a distinction should be made between two terms included in the entropy change dS (Fig. 1.26) (see *Kondepudi and Prigogine*, Further Reading). The first, d_eS, is the transfer of entropy *across system boundaries* due to energy and matter exchanges with the surroundings. In principle, these energy and matter exchanges can be performed reversibly [$\mathrm{d}_eS = \mathrm{d}Q_{rev}/T$; see the last equality of (1.111)], and the corresponding entropy change d_eS is either positive or negative (no thermodynamic law can be formulated on this term). In turn, the second term, d_iS, is the entropy produced *within* the system by irreversible processes that drive the system toward equilibrium. This term is always positive. The following expressions summarize these considerations:

$$\mathrm{d}S = \mathrm{d}_eS + \mathrm{d}_iS \quad \mathrm{d}_eS = \frac{\mathrm{d}Q}{T} \quad \mathrm{d}_iS > 0 \tag{1.115}$$

The statement according to which an irreversible process within the system has always a positive entropy change ($\mathrm{d}_iS > 0$) is one of the formulations of the second law of thermodynamics (**E20**). While the equilibrium is a time-independent state, an irreversible process has time as an associated external variable, and the arrow of time points toward equilibrium. Therefore, for an irreversible process,

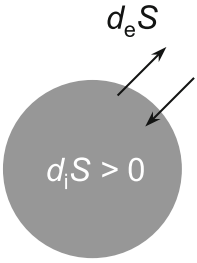


Fig. 1.26 Irreversible processes that occur inside the system are associated with an increase of entropy

$$\frac{\mathrm{d_i}S}{\mathrm{d}t} > 0 \tag{1.116}$$

where $\mathrm{d_i}S/\mathrm{d}t$ is called the **rate of entropy production**.

The first and second laws of thermodynamics enable the determination of energy and entropy changes, not of energy and entropy absolute values. However, in 1906, Walther Nernst (1864–1941) found that *the entropies of all systems approach zero as the temperature tends to zero*. This statement, usually called the third law of thermodynamics, has been empirically confirmed and enables the determination of absolute values for the entropy.

1.10 Irreversible Processes

We now consider examples of irreversible processes such as heat flow, gas expansion, matter diffusion, and chemical reactions. For each of them, we present the expressions of $\mathrm{d_i}S$ corresponding to a time interval $\mathrm{d}t$ and of the rate of entropy production $\mathrm{d_i}S/\mathrm{d}t$.

1.10.1 Heat Flow

We begin with an isolated system formed by two blocks of the same metal at two different temperatures T_1 and T_2 (Fig. 1.27). An isolated system does not exchange energy or matter with the surroundings, and so $\mathrm{d}V = 0$, $\mathrm{d}U = 0$, and $d_eS = 0$ (**E15**). For an infinitesimal heat $\mathrm{d}Q$ ($\mathrm{d}Q$ is an absolute value) passing from the metal block at higher temperature to the metal block at lower temperature, we can write

$$\mathrm{d_i}S = \frac{-\mathrm{d}Q}{T_1} + \frac{\mathrm{d}Q}{T_2} = \left(\frac{1}{T_2} - \frac{1}{T_1}\right)\mathrm{d}Q > 0 \tag{1.117}$$

since the entropy is an additive function, and for the rate of entropy production,

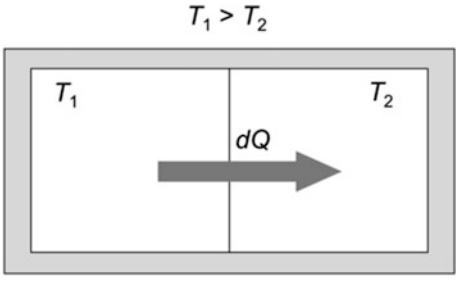


Fig. 1.27 In an isolated system, heat flows irreversibly from a higher to a lower temperature

$$\frac{\mathrm{d_i}S}{\mathrm{d}t} = \left(\frac{1}{T_2} - \frac{1}{T_1}\right)\frac{\mathrm{d}Q}{\mathrm{d}T} > 0 \tag{1.118}$$

with $\mathrm{d_i}S$ being produced in time $\mathrm{d}t$. As we know from experience, heat flows irreversibly from higher to lower temperature, and this process corresponds to $\mathrm{d_i}S$ greater than zero.

1.10.2 Gas Expansion

Consider now a closed system at a defined temperature ($\mathrm{d}T = 0$), with fixed boundaries ($\mathrm{d}V = 0$) and a frictionless wall separating a gas at pressures p_1, on one side of the piston, and p_2, on the other side of the piston ($p_1 > p_2$) (Fig. 1.28). The changes in volume of parts 1 and 2 of the system have the same absolute value, and so

$$\mathrm{d}V_1 = -\mathrm{d}V_2 \equiv \mathrm{d}V_\alpha \tag{1.119}$$

The system, consisting of a gas with a frictionless wall separating two parts at different pressures, exchanges heat with a thermostat in its surroundings, to maintain a constant temperature. The frictionless wall moves from higher to lower pressure, subject to the pressure difference $p_1 - p_2$ ($p_1 > p_2$). Then,

$$T\mathrm{d_e}S = \mathrm{d}U \tag{1.120}$$

and, making use of (1.119),

$$T\mathrm{d_i}S = p_1\mathrm{d}V_1 + p_2\mathrm{d}V_2 = (p_1 - p_2)\mathrm{d}V_\alpha > 0 \tag{1.121}$$

Therefore,

$$\frac{\mathrm{d_i}S}{\mathrm{d}t} = \left(\frac{p_1 - p_2}{T}\right)\frac{\mathrm{d}V_\alpha}{\mathrm{d}t} > 0 \tag{1.122}$$

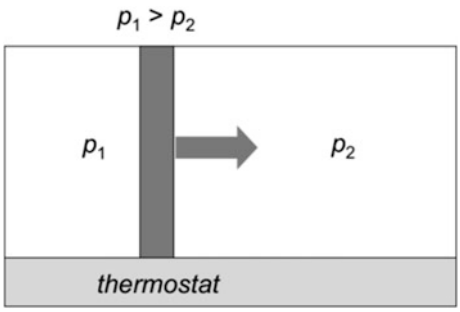


Fig. 1.28 At constant temperature, a frictionless wall moves from higher to lower pressure

1.10.3 Diffusion of Matter

The situation depicted in Fig. 1.29 represents a closed system consisting of two solutions with the same solute and same solvent, at constant temperature, separated by a semipermeable membrane that hypothetically allows the passage of the solute but not the solvent. It is further assumed that the parts of the system have the same volume, but the amount of solute is greater in part 1 than in part 2. Considering that the system has fixed boundaries ($dV = 0$), the changes in the amounts of solute are associated with irreversible diffusion of matter through the semipermeable membrane. The equation for dU is

$$dU = TdS + \mu_1 dn_1 + \mu_2 dn_2 \tag{1.123}$$

where

$$\mu_1 = \left(\frac{\partial U}{\partial n_1}\right)_{S,n_2} \quad \mu_2 = \left(\frac{\partial U}{\partial n_2}\right)_{S,n_1} \tag{1.124}$$

are called the **chemical potentials** for the solute in parts 1 and 2, and

$$-dn_1 = dn_2 \equiv d\xi \tag{1.125}$$

where dn_1 is negative (the amount of solute in part 1 decreases), dn_2 is positive (the amount of solute in part 2 increases) and $d\xi$ is the **extent** of solute spontaneously diffused through the semipermeable membrane in the time interval dt. From (1.124), we can conclude that the chemical potential is an intensive variable. The dn_1 and dn_2 terms account for changes in energy U due to solute transfer from part 1 to part 2. Expression (1.123) shows that $U = U(S, n_1, n_2)$, since the volume of the whole system is kept constant (V is not a variable in this experiment).

Making use of the first equality in (1.115), Eq. (1.123) can be split in two equations, one referring to the exchange of energy *across the system boundaries* (the whole system is kept at constant temperature),

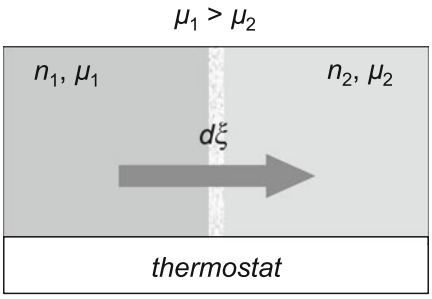


Fig. 1.29 At constant temperature, matter diffuses through the membrane from higher to lower chemical potential

$$\mathrm{d}U = T\mathrm{d}_{\mathrm{e}}S \tag{1.126}$$

the other being due to irreversible changes that occur *inside* the system, namely, in the amounts of the solute in parts 1 and 2, $\mathrm{d}n_1$ and $\mathrm{d}n_2$, and in the corresponding entropy change $\mathrm{d}_{\mathrm{i}}S$,

$$0 = T\mathrm{d}_{\mathrm{i}}S + \mu_1\mathrm{d}n_1 + \mu_2\mathrm{d}n_2 \tag{1.127}$$

Note that addition of (1.126) and (1.127) gives back (1.123). Substitution of (1.125) in (1.127) leads to

$$d_iS = -\frac{1}{T}(\mu_1 dn_1 + \mu_2 dn_2) = \frac{\mu_1 - \mu_2}{T}d\xi > 0 \tag{1.128}$$

and

$$\frac{\mathrm{d}_{\mathrm{i}}S}{\mathrm{d}t} = \left(\frac{\mu_1 - \mu_2}{T}\right)\frac{\mathrm{d}\xi}{\mathrm{d}t} > 0 \tag{1.129}$$

Therefore, **diffusion of matter** occurs from the higher chemical potential μ_1 to the lower chemical potential μ_2. In the irreversible process of matter flow, μ_1 decreases, whereas μ_2 increases until equality between the chemical potentials μ_1 and μ_2 is attained at equilibrium.

1.10.4 Chemical Reaction

Consider now a closed system with the following **chemical reaction** at a defined temperature T and pressure p,

$$a\mathrm{A(g)} + b\mathrm{B(g)} \rightarrow c\mathrm{C(g)} + d\mathrm{D(g)} \tag{1.130}$$

where a, b, c, and d are the **stoichiometric coefficients**. Infinitesimal changes in the amounts of reactants A and B and products C and D are related to one another by the reaction stoichiometry and

$$\frac{\mathrm{d}n_{\mathrm{A}}}{-a} = \frac{\mathrm{d}n_{\mathrm{B}}}{-b} = \frac{\mathrm{d}n_{\mathrm{C}}}{c} = \frac{\mathrm{d}n_{\mathrm{D}}}{d} \equiv d\xi \tag{1.131}$$

where the identity defines the state variable ξ, called the **extent** of the chemical reaction. The stoichiometric coefficients for the reactants are multiplied by -1 because the infinitesimal changes in the corresponding amounts of the reactants, $\mathrm{d}n_{\mathrm{A}}$ and $\mathrm{d}n_{\mathrm{B}}$, are negative. Equation (1.131) can be rewritten in the more general form

$$\mathrm{d}n_k = \nu_k\mathrm{d}\xi \tag{1.132}$$

where ν_k takes the values $-a$, $-b$, c, d, which are called **stoichiometric numbers**. Combining the first law (1.87) and the second law (1.115) and generalizing from (1.127), we can write

$$\mathrm{d}U = T\mathrm{d}_\mathrm{e}S - p\mathrm{d}V + T\mathrm{d}_\mathrm{i}S + \mu_\mathrm{A}\mathrm{d}n_\mathrm{A} + \mu_\mathrm{B}\mathrm{d}n_\mathrm{B} + \mu_\mathrm{C}\mathrm{d}n_\mathrm{C} + \mu_\mathrm{D}\mathrm{d}n_\mathrm{D} \tag{1.133}$$

and split this equation into

$$\mathrm{d}U = T\mathrm{d}_\mathrm{e}S - p\mathrm{d}V \quad 0 = T\mathrm{d}_\mathrm{i}S + \mu_\mathrm{A}\mathrm{d}n_\mathrm{A} + \mu_\mathrm{B}\mathrm{d}n_\mathrm{B} + \mu_\mathrm{C}\mathrm{d}n_\mathrm{C} + \mu_\mathrm{D}\mathrm{d}n_\mathrm{D} \tag{1.134}$$

In turn, substitution of (1.131) in the second equation of (1.134) leads to

$$\begin{aligned} \mathrm{d}_\mathrm{i}S &= -\frac{1}{T}(\mu_\mathrm{A}\mathrm{d}n_\mathrm{A} + \mu_\mathrm{B}\mathrm{d}n_\mathrm{B} + \mu_\mathrm{C}\mathrm{d}n_\mathrm{C} + \mu_\mathrm{D}\mathrm{d}n_\mathrm{D}) \\ &= -\frac{1}{T}(-a\mu_\mathrm{A} - b\mu_\mathrm{B} + c\mu_\mathrm{C} + \mathrm{d}\mu_\mathrm{D})\mathrm{d}\xi \end{aligned} \tag{1.135}$$

Theophile De Donder (1872–1957) defined the state variable called **affinity** of the chemical reaction as the stoichiometric sum of the chemical potentials of the reactants minus the stoichiometric sum of the chemical potentials of the products,

$$A = a\mu_\mathrm{A} + b\mu_\mathrm{B} - c\mu_\mathrm{C} - d\mu_\mathrm{D} \tag{1.136}$$

(see *Kondepudi*, Further Reading). Substitution of (1.136) in (1.135) gives

$$\mathrm{d}_\mathrm{i}S = \frac{A}{T}\mathrm{d}\xi > 0 \tag{1.137}$$

Considering now that $\mathrm{d}\xi$ refers to the time interval $\mathrm{d}t$, we can write

$$\frac{\mathrm{d}_\mathrm{i}S}{\mathrm{d}t} = \left(\frac{A}{T}\right)\frac{\mathrm{d}\xi}{\mathrm{d}t} > 0 \tag{1.138}$$

where $\mathrm{d}\xi/\mathrm{d}t$ is the **rate of the chemical reaction**. As the chemical reaction proceeds to equilibrium, where $\mathrm{d}\xi/\mathrm{d}t = 0$, the chemical affinity decreases to zero, a value attained at equilibrium.

1.11 Chemical Potential

The concept of chemical potential is essential to describing the change of composition of a mixture, be it a physical change or a chemical reaction. For that reason, it is important to find a functional dependence between chemical potential and pressure of gases and between chemical potential and concentration in solutions.

1.11.1 Gibbs–Duhem Equation

The functional dependence between chemical potential and pressure uses an important thermodynamic equation known as the **Gibbs–Duhem equation** [Josiah Willard Gibbs (1839–1903), Pierre Duhem (1861–1916)], which shows that the intensive variables temperature, pressure, and chemical potential are not independent.

We begin by combining the first and second laws for a closed system with possible changes in the amounts n_k of the system components due to a chemical reaction and write

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V + \sum_k \mu_k \mathrm{d}n_k \tag{1.139}$$

From this equation, we conclude that

$$\left(\frac{\partial U}{\partial S}\right)_{V,n_k} = T \quad \left(\frac{\partial U}{\partial V}\right)_{S,n_k} = -p \quad \left(\frac{\partial U}{\partial n_k}\right)_{S,V,n_j(j\neq k)} = \mu_k \tag{1.140}$$

From (1.139), we learn that the energy U is a function of entropy S, volume V, and amount of substance n_k for all k values, $U = U(S, V, n_1, \ldots, n_k, \ldots)$. Here U, S, V, and n_k are extensive variables. Hence, multiplication of S, V, and all the n_k by a factor λ is equivalent to multiplying U by the same factor,

$$U(\lambda S, \lambda V, \lambda n_1, \ldots, \lambda n_k, \ldots) = \lambda U(S, V, n_1, \ldots, n_k, \ldots) \tag{1.141}$$

Differentiating both members of (1.141) with respect to λ (differentiation of the first member requires the chain rule for partial differentiation) leads to

$$\frac{\partial U}{\partial(\lambda S)}\frac{\partial(\lambda S)}{\partial \lambda} + \frac{\partial U}{\partial(\lambda V)}\frac{\partial(\lambda V)}{\partial \lambda} + \sum_k \frac{\partial U}{\partial(\lambda n_k)}\frac{\partial(\lambda n_k)}{\partial \lambda} = \frac{\mathrm{d}(\lambda U)}{\mathrm{d}\lambda}$$

which simplifies to

$$\frac{\partial U}{\partial(\lambda S)}S + \frac{\partial U}{\partial(\lambda V)}V + \sum_k \frac{\partial U}{\partial(\lambda n_k)}n_k = U$$

Considering now $\lambda = 1$ (λ is an arbitrary scaling factor), we obtain

$$\frac{\partial U}{\partial S}S + \frac{\partial U}{\partial V}V + \sum_k \frac{\partial U}{\partial n_k}n_k = U \tag{1.142}$$

Substitution of (1.140) leads to

$$U = TS - pV + \sum_k \mu_k n_k \tag{1.143}$$

In turn, differentiation of (1.143) gives

$$\mathrm{d}U = T\mathrm{d}S + S\mathrm{d}T - p\mathrm{d}V - V\mathrm{d}p + \sum_k \mu_k \mathrm{d}n_k + \sum_k n_k \mathrm{d}\mu_k \tag{1.144}$$

If we now subtract (1.139) from (1.144), we obtain

$$0 = S\mathrm{d}T - V\mathrm{d}p + \sum_k n_k \mathrm{d}\mu_k \tag{1.145}$$

This equality, known as the **Gibbs–Duhem equation**, shows that the intensive variables T, p, and μ_k are not independent variables. For a system with one component, the Gibbs–Duhem equation takes the form

$$\mathrm{d}\mu = -S_\mathrm{m}\mathrm{d}T + V_\mathrm{m}\mathrm{d}p \tag{1.146}$$

where S_m and V_m represent the molar entropy and molar volume, respectively. At a specified temperature ($\mathrm{d}T = 0$), integration of (1.146) gives

$$\int_{\mu^0}^{\mu} d\mu = \int_{p^0}^{p} V_m dp \tag{1.147}$$

where the lower limits are the chemical potential and pressure of a reference state (usually, $p^0 = 1$ bar), and the upper limits are the chemical potential and pressure as variables (this equation is an equality between indefinite integrals). After integration of the first member, we obtain

$$\mu(p,T) = \mu^0(T) + \int_{p^0}^{p} V_m(p,T)\mathrm{d}p \tag{1.148}$$

We will now proceed by applying this equation to ideal and real gases, liquid solutions, pure liquids, and solids.

1.11.2 Ideal Gas

For an ideal gas, $V_\mathrm{m} = RT/p$. Hence,

$$\mu(p,T) = \mu^0(T) + RT\ \ln\left(\frac{p}{p^0}\right) \tag{1.149}$$

Considering an ideal gas mixture, the chemical potential for the k component is given by

$$\mu_k(p_k, T) = \mu_k^0(T) + RT \ln\left(\frac{p_k}{p^0}\right) \tag{1.150}$$

where the partial pressure for the k component is given by

$$p_k = x_k p \tag{1.151}$$

with x_k the molar fraction for the k component and p the total pressure of the gaseous mixture. Substitution of (1.151) in (1.150) gives

$$\mu_k(p, T, x_k) = \left[\mu_k^0(T) + RT \ln\left(\frac{p}{p^0}\right)\right] + RT \ln x_k = \mu_k^0(p, T) + RT \ln x_k \tag{1.152}$$

1.11.3 Real Gases

Let us now consider the chemical potential for a real gas and its deviation from the corresponding ideal gas that defines the reference state. In order to compare molar volumes of real and ideal gases, we take the **compressibility factor** Z defined by

$$V_{\mathrm{m,real}} = ZV_{\mathrm{m,ideal}} = Z\frac{RT}{p} \tag{1.153}$$

This definition shows that $Z = 1$ for the ideal gas. Subtracting $V_{\mathrm{m,ideal}}$ from each member of the above equation leads to

$$V_{\mathrm{m,real}} - V_{\mathrm{m,ideal}} = (Z-1)V_{\mathrm{m,ideal}} = RT\frac{Z-1}{p} \tag{1.154}$$

Making use of (1.148) for the real and ideal gases and noting that the standard state is defined by the ideal gas, we can write

$$\begin{aligned} \mu_{\mathrm{real}}(p, T) &= \mu_{\mathrm{ideal}}^0(T) + \int_{p^0}^{p} V_{\mathrm{m,real}}(p, T)\mathrm{d}p \\ \mu_{\mathrm{ideal}}(p, T) &= \mu_{\mathrm{ideal}}^0(T) + \int_{p^0}^{p} V_{\mathrm{m,ideal}}(p, T)\mathrm{d}p \end{aligned} \tag{1.155}$$

Now we take the difference between these equations and use (1.154) to obtain

$$\mu_{\text{real}}(p,T) = \mu_{\text{ideal}}(p,T) + RT \int_{p^0}^{p} \left(\frac{Z-1}{p}\right) dp \tag{1.156}$$

G.N. Lewis (1875–1946) defined a quantity with pressure units called **fugacity** f that gives a pressure-dependence similar to that of the ideal gas,

$$\mu_{\text{real}}(p,T) = \mu_{\text{ideal}}(p,T) + RT \ln\left(\frac{f}{p}\right) \tag{1.157}$$

[compare with (1.149)]. Fugacity is a corrected pressure that takes into account deviations from the ideal gas's behavior due to the presence of intermolecular interactions. When the pressure approaches zero, a real gas acquires ideal gas behavior. Hence,

$$\lim_{p\to 0}\left(\frac{f}{p}\right) = 1 \tag{1.158}$$

1.11.4 Liquid Solutions

Diluted nonionic solutions exhibit a chemical potential-dependence on the solute mole fraction similar to that of ideal solutions [see the last equation of (1.152)],

$$\mu_k(T,x_k) = \mu_k^0(T) + RT \ln x_k \tag{1.159}$$

where $x_k \ll 1$ and the dependence on pressure has been ignored, since liquids, and condensed phases (liquids and solids) in general, have little sensitivity to pressure. For nonideal solutions, Lewis introduced the concept of **activity** to keep the same form of (1.159), and wrote

$$\mu_k(T,x_k) = \mu_k^0(T) + RT \ln a_k \tag{1.160}$$

where the activity a_k is given by

$$a_k = \gamma_k x_k \tag{1.161}$$

and γ_k is called the **activity coefficient** ($\gamma_k = 1$ implies $a_k = x_k$, that is, ideal behavior).

1.11.5 Pure Liquids and Solids

Returning to the integration of both members of (1.146), between 0 K and temperature T, and pressure between p^0 and p, we obtain

$$\mu(p,T) = \mu(p^0,0) - \int_0^T S_m(T)\mathrm{d}T + \int_{p^0}^{p} V_\mathrm{m}\mathrm{d}p \tag{1.162}$$

Solids and liquids are little affected by pressure changes. Hence, the molar volume can be considered approximately a constant. Then,

$$\mu(p,T) = \mu^0(T) + V_\mathrm{m}(p - p^0) \tag{1.163}$$

where

$$\mu^0(T) \equiv \mu(p^0,0) - \int_0^T S_\mathrm{m}(T)\mathrm{d}T \tag{1.164}$$

Making use of (1.160) for one component and of (1.163), we can write

$$\mu^0(T) + V_\mathrm{m}(p - p^0) = \mu^0(T) + RT\ln a \tag{1.165}$$

For solids and liquids, $\mu^0(T)$ is several orders of magnitude greater than $V_\mathrm{m}\Delta p$. For instance, for liquid water and Δp = 1 bar (= 100000 Pa), $V_\mathrm{m}\Delta p$ = 1.8 J mol^{-1} and $\mu^0(T)$ = −237.1 kJ mol^{-1} (chemical potential value taken from *Handbook of Chemistry and Physics*, 2011). Hence, within a good approximation, the activity of pure solids and liquids is equal to 1 ($\ln a \approx 0$).

1.12 Gibbs Energy

At constant temperature and pressure, the state function that evolves to a minimum at equilibrium is the **Gibbs energy**, defined by the following identity:

$$G \equiv U + pV - TS \tag{1.166}$$

Differentiating G and using (1.89) and (1.115) for a closed system, we obtain

$$\mathrm{d}G = \mathrm{d}U + p\mathrm{d}V + V\mathrm{d}p - T\mathrm{d}_\mathrm{e}S - T\mathrm{d}_\mathrm{i}S - S\mathrm{d}T = -S\mathrm{d}T + V\mathrm{d}p - T\mathrm{d}_\mathrm{i}S \tag{1.167}$$

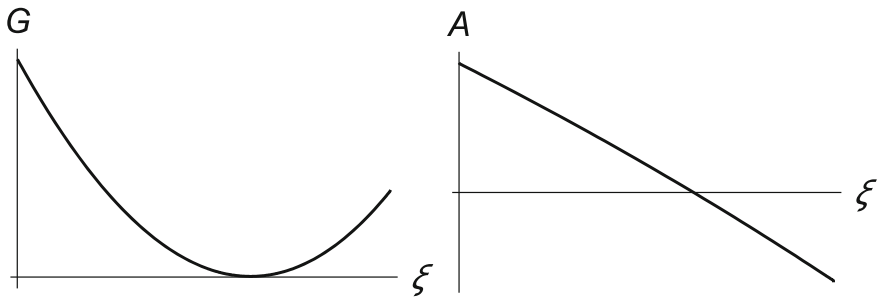


Fig. 1.30 At equilibrium, the Gibbs energy attains a minimum, and the reaction affinity is zero. Graphs obtained with *Mathematica*

At constant temperature and pressure ($dT = 0$, $dp = 0$), we can write

$$dG_{T,p} = -Td_iS \leq 0 \tag{1.168}$$

This result shows that at constant temperature and pressure, the Gibbs energy evolves through an irreversible process to a minimum at equilibrium ($d_iS > 0$, and so $dG_{T,p}$ is negative; at the minimum, the entropy is a maximum). When the irreversible process is a chemical reaction,

$$dG_{T,p} = -T d_iS = \sum_k \mu_k dn_k = -A d\xi \tag{1.169}$$

where A is the chemical reaction affinity [see (1.132) and (1.136)]. Making use of (1.63) in (1.169), we can write

$$\left(\frac{\partial G}{\partial n_k}\right)_{T,p,n_j(j\neq k)} = \mu_k \quad \left(\frac{\partial G}{\partial \xi}\right)_{T,p} = -A \tag{1.170}$$

The second of the above equations shows that the partial derivative of the Gibbs energy with respect to the extent of the reaction, at constant temperature and constant pressure, is the negative of the chemical reaction affinity. At equilibrium, the Gibbs energy attains a minimum, and the chemical reaction affinity is zero (Fig. 1.30).

1.12.1 Chemical Potential and Gibbs Energy of Formation

The first of the above equations shows that the partial derivative of the Gibbs energy with respect to the amount of substance of the k component, at constant temperature, constant pressure, and fixed amounts of other substances in the reaction mixture, is equal to the chemical potential of the k component of the

chemical reaction mixture. While this is the formal relationship between chemical potential and Gibbs energy (the chemical potential is the partial molar Gibbs energy), we can proceed to find a more pragmatic expression between chemical potential and Gibbs energy as experimentally measurable physical quantities.

Gibbs energy is an extensive state variable. In order to understand the relationship between standard Gibbs energies of formation and standard chemical potentials, we substitute (1.143), obtained during the derivation of the Gibbs–Duhem equation, into (1.166), the definition of Gibbs energy, leading to

$$G = \sum_k \mu_k n_k \tag{1.171}$$

For one compound, we can write

$$\mu = G_\mathrm{m} = \frac{G}{n} \tag{1.172}$$

thus concluding that the chemical potential is a molar Gibbs energy, G_m (chemical potential and molar Gibbs energy have the same units, energy per mole). Therefore, G is an extensive state function proportional to n, with the proportionality constant being the chemical potential. However, in order to determine Gibbs energies experimentally, zeros for the Gibbs energies and chemical potentials need to be established. As mentioned before, the **standard state** of a pure substance at a specified temperature is its most stable state (gas, liquid, or pure crystalline solid) at the standard pressure $p^0 = 1$ bar. At a specified temperature, the **standard molar Gibbs energy of formation** ΔG_f^0 of a particular compound is the difference between the molar Gibbs energy of this compound in its standard state and the Gibbs energies of its constituent elements in their standard states. Since atoms of elemental substances do not transmute in chemical reactions (only nuclear reactions outside the thermodynamic realm can lead to transmutation of chemical elements), *all elemental substances* in their standard states are assumed to have a standard chemical potential and a standard molar Gibbs energy of formation equal to zero *at all temperatures*,

$$\mu_\alpha^0 = \Delta G_{\mathrm{f},\alpha}^0 = 0 \tag{1.173}$$

where α generically denotes elemental substances. Having established the zeros for the standard chemical potentials and the standard molar Gibbs energies of formation of compounds, we can now conclude that at a specified temperature, the standard chemical potential of substance k is equal to the standard Gibbs energy of formation of the same substance,

$$\mu_\alpha^0 = \Delta G_{\mathrm{f},k}^0 \tag{1.174}$$

where k stands for a pure compound. At a specified temperature, the **standard Gibbs energy of a chemical reaction**, ΔG_r^0, is the difference between the stoichiometric sum of the standard Gibbs energies of formation of the products and the stoichiometric sum of the standard Gibbs energies of formation of the reactants, with reactants and products being in their standard states.

1.12.2 Gibbs–Helmholtz Equation

The Gibbs–Helmholtz equation relates the temperature-dependence of the Gibbs energy to the corresponding enthalpy value and can be used to determine the enthalpy variation from data on the temperature variation of the Gibbs energy. To arrive at the **Gibbs–Helmholtz equation**, we begin by considering a closed system at equilibrium ($d_i S = 0$) and substitute this equation into (1.167), thus obtaining

$$\left(\frac{\partial G}{\partial T}\right)_p = -S \tag{1.175}$$

Combining the definition of Gibbs energy [see (1.166) and (1.92)] with (1.175), we can write

$$G = H + T\left(\frac{\partial G}{\partial T}\right)_p \tag{1.176}$$

By dividing both members of this equality by T^2 and rearranging, we obtain

$$\frac{1}{T}\left(\frac{\partial G}{\partial T}\right)_p - \frac{G}{T^2} = -\frac{H}{T^2} \tag{1.177}$$

that is,

$$\frac{\partial\left(\frac{1}{T}G\right)}{\partial T} = -\frac{H}{T^2} \tag{1.178}$$

If we now apply this equation to the Gibbs energy variation of a chemical reaction ΔG and the corresponding enthalpy change ΔH, we obtain

$$\frac{\partial\left(\frac{1}{T}\Delta G\right)}{\partial T} = -\frac{\Delta H}{T^2} \tag{1.179}$$

1.13 Chemical Equilibrium

Consider now the chemical reaction

$$a\mathrm{A(g)} + b\mathrm{B(g)} \rightarrow c\mathrm{C(g)} \; + \; d\mathrm{D(g)} \tag{1.180}$$

where both reactants and products are gases. Breaking down the Gibbs energy of the chemical reaction into the Gibbs energies of formation of products and reactants, and using the relationship between Gibbs energy of formation and chemical potential, we can write

$$\Delta G_\mathrm{r} = [(c\Delta G_{\mathrm{f,C}} + d\Delta G_{\mathrm{f,D}}) - (a\Delta G_{\mathrm{f,A}} + b\Delta G_{\mathrm{f,B}})] = (c\mu_\mathrm{C} + d\mu_\mathrm{D}) - (a\mu_\mathrm{A} + b\mu_\mathrm{B}) \tag{1.181}$$

Use of (1.160) allows writing

$$\Delta G_\mathrm{r} = \left[\left(c\mu_\mathrm{C}^0 + d\mu_\mathrm{D}^0\right) - \left(a\mu_\mathrm{A}^0 + b\mu_\mathrm{B}^0\right)\right] + RT\,\ln\chi \tag{1.182}$$

where the **reaction quotient** χ is given by

$$\chi = \frac{a_\mathrm{C}^c a_\mathrm{D}^d}{a_\mathrm{A}^a a_\mathrm{B}^b} \tag{1.183}$$

An equation equivalent to (1.181) can be written for the standard states of products and reactants

$$\Delta G_\mathrm{r}^0 = [(c\Delta G_{\mathrm{f,C}}^0 + d\Delta G_{\mathrm{f,D}}^0) - (a\Delta G_{\mathrm{f,A}}^0 + b\Delta G_{\mathrm{f,B}}^0)] = (c\mu_\mathrm{C}^0 + d\mu_\mathrm{D}^0) - (a\mu_\mathrm{A}^0 + b\mu_\mathrm{B}^0) \tag{1.184}$$

Combining (1.182) and (1.184) leads to

$$\Delta G_\mathrm{r} = \Delta G_\mathrm{r}^0 + RT\,\ln\chi \tag{1.185}$$

When approaching chemical equilibrium, ΔG_r tends to zero, and the reaction quotient tends to the **equilibrium constant**,

$$\Delta G_\mathrm{r} \rightarrow 0 \quad \chi \rightarrow K_\mathrm{e} \tag{1.186}$$

where

$$K_\mathrm{e} = \frac{a_{\mathrm{C,eq}}^c a_{\mathrm{D,eq}}^d}{a_{\mathrm{A,eq}}^a a_{\mathrm{B,eq}}^b} \tag{1.187}$$

Applying (1.186) to (1.185) gives

$$\Delta G_r^0 = -RT \ln K_e \tag{1.188}$$

(**E16**; see *Mathematica* code **M7**). From (1.160) and (1.184), it can be concluded that the equilibrium constant is a function of temperature only. This important result is at the origin of the expression "equilibrium constant."

1.14 Gibbs Phase Rule

In thermodynamics, distinct physical states of matter are called **phases**. For a one-component system, the occurrence of different phases and the equilibria between them can be represented in a graph of pressure as a function of temperature, called a **phase diagram**. The p–T curves of a typical phase diagram for a one-component system (Fig. 1.31) indicate two-phase equilibria, namely liquid $\rightleftarrows$ gas, solid $\rightleftarrows$ liquid, and solid $\rightleftarrows$ gas equilibria. The point t that gives the pressure and temperature of the three-phase equilibrium solid $\rightleftarrows$ liquid $\rightleftarrows$ gas is called the **triple point**. The point c is the **critical point** T_c, the temperature above which a supercritical fluid cannot be liquefied by increase of pressure.

Let us consider a closed system consisting of two phases, liquid and gas (Fig. 1.31). At equilibrium, the chemical potentials of the liquid and gas are equal,

$$\mu_{liq}(p,T) = \mu_{gas}(p,T) \tag{1.189}$$

there are no irreversible processes in the system, and the entropy production is zero [see (1.169)],

$$\frac{d_i S}{dt} = \left[\frac{\mu_{liq}(p,T) - \mu_{gas}(p,T)}{T}\right]\frac{d\xi}{dt} = 0 \tag{1.190}$$

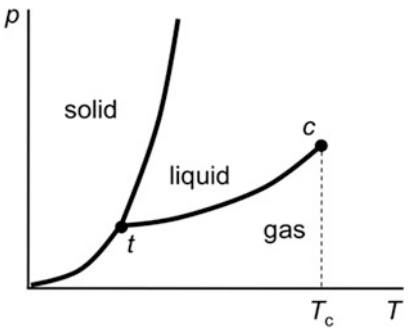


Fig. 1.31 Typical phase diagram for a one-component system

Hence,

$$\mu^0_{\text{liq}}(T) = \mu^0_{\text{gas}}(T) + RT \ln\left(\frac{p}{p^0}\right) \tag{1.191}$$

[see (1.149) and (1.163)], where p is the vapor pressure of the pure liquid. After rearrangement, we can write

$$\ln\left(\frac{p}{p^0}\right) = \frac{\mu^0_{\text{liq}}(T) - \mu^0_{\text{gas}}(T)}{RT} \tag{1.192}$$

(**E17**). This equation allows us to conclude that the vapor pressure is a function of temperature, a functional dependence that corresponds to the p–T curve of the liquid–gas equilibrium in the phase diagram. Equilibria between any two phases of one-component systems also lead to p-T curves of phase diagrams. Note that these curves relate two intensive variables, namely p and T.

Besides pressure and temperature, solutions have additional intensive variables related to composition. Let us consider a system at equilibrium containing C components in each of its P phases: how many independent intensive variables or **degrees of freedom** does such a system have?

If composition variables are expressed by mole fractions, then the sum of all mole fractions in each phase equals 1. In particular, for phase α,

$$\sum_{k=1}^{C} x_k^{(\alpha)} = 1 \tag{1.193}$$

Therefore, we have $C - 1$ independent mole fractions in each phase and $P(C - 1)$ independent mole fractions altogether, to which we should add two additional intensive variables corresponding to pressure and temperature. Then the total number of intensive variables in the system is equal to $P(C - 1) + 2$. For a particular component, say component k, the equations involving P phases are given by

$$\mu_k^1(p,T) = \mu_k^2(p,T) = \ldots = \mu_k^P(p,T) \tag{1.194}$$

and their number is equal to $P - 1$. For C system **components**, we have $C(P - 1)$ equations. Hence, the number of independent intensive variables in the system, that is, the number of degrees of freedom f, is equal to the number of intensive variables $P(C - 1) + 2$ minus the number of equations involving those variables $C(P - 1)$,

$$f = P(C-1) + 2 - C(P-1) = C - P + 2 \tag{1.195}$$

This result is known as the **Gibbs phase rule**. For $C = 1$ and $P = 1$ (one-component system and one phase), $f = 2$. For $C = 1$ and $P = 2$ (one-component system and two phases in equilibrium), $f = 1$. Finally, for $C = 1$

and $P = 3$ (one-component system and three phases in equilibrium), $f = 0$ (zero degrees of freedom corresponds to the triple point).

When the system components are involved in R chemical equilibria, there are R additional equations expressing the affinities of the chemical reactions,

$$A_1 = 0 \quad A_2 = 0 \quad \ldots \quad A_R = 0 \tag{1.196}$$

Each of these equations makes each chemical potential dependent on the remaining ones, thus reducing by 1 the number of independent chemical potentials [see the first equation of (1.170)]. For R chemical equilibria, the number of **independent components** is reduced by R. Therefore, the Gibbs phase rule takes the following more general form:

$$f = (C - R) - P + 2 = C_\mathrm{i} - P + 2 \tag{1.197}$$

where the number of independent components C_i is obtained by subtracting the number of chemical equilibria R from the number of components C in the system (**E18**).

1.15 Helmholtz Energy

At constant temperature and volume, the state function that evolves to a minimum at equilibrium is the **Helmholtz energy**, defined by the following identity:

$$F \equiv U - TS \tag{1.198}$$

Differentiating (1.198) and using (1.89) and (1.115), we obtain

$$\mathrm{d}F = \mathrm{d}U - T\mathrm{d}S - S\mathrm{d}T = \mathrm{d}U - T\mathrm{d}_\mathrm{e}S - T\mathrm{d}_\mathrm{i}S - S\mathrm{d}T = -p\mathrm{d}V - S\mathrm{d}T - T\mathrm{d}_\mathrm{i}S \tag{1.199}$$

for a closed system. At constant temperature and volume ($\mathrm{d}T = 0$, $\mathrm{d}V = 0$),

$$\mathrm{d}F_{T,V} = -T\mathrm{d}_\mathrm{i}S \leq 0 \tag{1.200}$$

where the inequality results from the second law for irreversible processes ($\mathrm{d}_\mathrm{i}S > 0$), and the equality corresponds to equilibrium ($\mathrm{d}_\mathrm{i}S = 0$). Therefore, in a closed system at constant temperature and volume, the Helmholtz energy evolves through an irreversible process to a minimum at equilibrium ($\mathrm{d}F_{T,V}$ is negative). When the irreversible process is a chemical reaction,

$$\mathrm{d}F_{T,V} = -T\mathrm{d}_\mathrm{i}S = \sum_k \mu_k \mathrm{d}n_k = -A\mathrm{d}\xi \tag{1.201}$$

where A is the chemical reaction affinity [see (1.132) and (1.136)]. Making use of (1.63) in (1.201), we can write

$$\left(\frac{\partial F}{\partial n_k}\right)_{T,V,n_j(j\neq k)} = \mu_k \quad \left(\frac{\partial F}{\partial \xi}\right)_{T,V} = -A \tag{1.202}$$

The second of the above equations shows that the partial derivative of the Helmholtz energy with respect to the extent of the reaction, at constant temperature and volume, is the negative of the chemical reaction affinity. At equilibrium, the Helmholtz energy attains a minimum, and the chemical reaction affinity is zero.

1.16 Surface Tension

1.16.1 Liquid Droplet in Air

Consider now a system formed by a small liquid droplet in the air of a box, at constant temperature and volume (Fig. 1.32). Since the gravitational attraction on the droplet is negligible, the droplet takes an approximately spherical shape. It is found that the diameter of the droplet is smaller than that of a sphere with the same mass in the liquid bulk. At constant temperature and volume, the liquid droplet reduces its diameter as the contracting force on its surface due to the **surface tension** is balanced by an increase in the inside pressure (Fig. 1.32).

Let p' and V' be the pressure and air volume in the box, and p'' and V'' the pressure and volume of the droplet. Then, using (1.87) and the second equation of (1.115), we can write

$$\mathrm{d}U = T\mathrm{d}S - p'\mathrm{d}V' - p''\mathrm{d}V'' + \gamma\mathrm{d}\Sigma \tag{1.203}$$

The last term of this expression is the work required to increase the interfacial area by $\mathrm{d}\Sigma$, and the proportionality coefficient γ is called **surface tension**, whose units are energy per area or force per length. Substitution of (1.203) in the first equality of (1.199) leads to

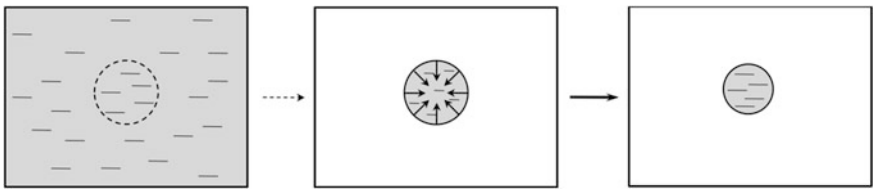


Fig. 1.32 At constant temperature and volume, a suspended liquid droplet reduces its diameter in an irreversible process that ends when the contracting force of the surface is balanced by the increase in the inside pressure. In this irreversible process, the Helmholtz energy evolves to a minimum

$$dF = -p'dV' - p''dV'' - SdT + \gamma d\Sigma \tag{1.204}$$

where

$$\gamma = \left(\frac{\partial F}{\partial \Sigma}\right)_{T,V',V''} \tag{1.205}$$

that is, the surface tension is the Helmholtz energy per unit of interfacial area, at constant T, V', and V''. At 25 °C, water has a surface tension equal to 71.99 mJ m^{-2}, the surface tension of mercury is 485.48 mJ m^{-2}, and most organic compounds have, at temperatures below their boiling points, surface tensions between 15 and 50 mJ m^{-2} (values taken from *Handbook of Chemistry and Physics*, 2011).

The contraction of the liquid droplet in air is an irreversible process that makes the system evolve to equilibrium in a time interval dt. Introducing constant temperature ($dT = 0$) and constant overall volume ($dV' + dV'' = 0$) in (1.204) and making use of (1.200), we obtain

$$-T\frac{d_i S}{dt} = -(p'' - p')\frac{dV''}{dt} + \gamma\frac{d\Sigma}{dt} \tag{1.206}$$

where V'' is the volume of the liquid droplet ($V'' = 4\pi r^3/3$, $dV'' = 4\pi r^2 dr$) and Σ is its surface area ($\Sigma = 4\pi r^2$, $d\Sigma = 8\pi r dr$). Substitution of these values in (1.206) leads to

$$-T\frac{d_i S}{dt} = \left[-(p'' - p')4\pi r^2 + \gamma 8\pi r\right]\frac{dr}{dt} \tag{1.207}$$

During the evolution to equilibrium, the expression in square brackets tends to zero and impels dr/dt to tend to zero, in the same way as a temperature gradient is the cause for the corresponding irreversible heat flow or a concentration gradient is the cause for an irreversible flow of matter. Therefore, at equilibrium, both the expression in square brackets and dr/dt are zero. Hence, the **excess pressure** in the liquid droplet can be expressed in terms of the surface tension and the droplet radius,

$$\Delta p = p'' - p' = 2\gamma/r \tag{1.208}$$

1.16.2 Capillary Action

When the lower end of a vertical capillary tube is dipped into water, water rises in the capillary tube as a manifestation of surface tension. Unlike the liquid droplet in air, where there was a single interfacial surface, the liquid–air interface, the capillary tube dipped into water has two interfaces involving water, namely the water–air and water–glass interfaces. Water wets the glass, and so water rises in the capillary tube and the meniscus is concave.

The intersection between the liquid–air interface and the liquid–solid interface is a circumference of a circle whose radius is equal to the internal radius of the capillary tube. The angle at which the liquid is in contact with the capillary tube is called the **contact angle** and is denoted by θ (Fig. 1.33a). If the liquid wets the glass, then this angle is smaller than $\pi/2$ ($\cos\theta > 0$) and the height of liquid inside the capillary tube is denoted by h ($h > 0$) (Fig. 1.33a). Since temperature is kept constant and the total volume of liquid is also constant, the thermodynamic state function that evolves to a minimum (equilibrium) under these conditions is the Helmholtz energy [see the equality in (1.200)]. In the irreversible process leading to equilibrium, the liquid inside the capillary tube gradually rises, and the pressure ρgh times the volume of the rising column of liquid inside the capillary tube (in units, pressure $\times$ volume = energy) is positive and increases. In turn, the surface tension $\gamma \cos\theta$ times the area of the interior wall of the capillary tube (in units, surface tension $\times$ area = energy) is negative and decreases as h increases. Hence, the minimum of the Helmholtz energy results from a balance between these two opposing quantities and is given by

$$\begin{aligned} \Delta F(h) &= -(\gamma \cos\theta) 2\pi r h + \int_0^h (\rho h g) \pi r^2 \mathrm{d}h \\ &= -(\gamma \cos\theta) 2\pi r h + \frac{\rho g \pi r^2 h^2}{2} \end{aligned} \tag{1.209}$$

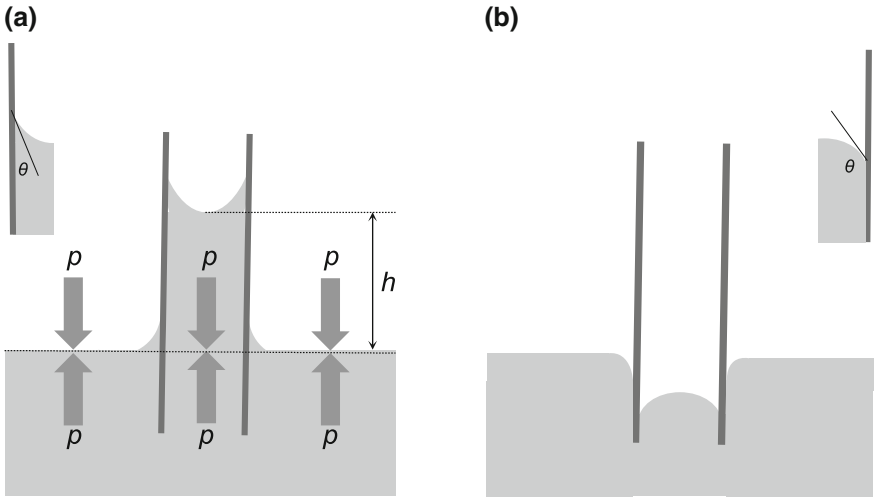


Fig. 1.33 **a** For a liquid that wets the glass, one observes capillary rise, and the meniscus is concave. The contact angle θ is less than $\pi/2$ and $\cos\theta > 0$. At equal heights throughout the liquid, pressure is equal. **b** For a contact angle θ greater than $\pi/2$, $\cos\theta < 0$, and the liquid is depressed in the capillary tube. This is the case of mercury, for example

Minimization of ΔF as a function of h leads to the value of h at equilibrium, that is,

$$\frac{\partial \Delta F(h)}{\partial h} = -(\gamma \cos\theta)2\pi r + \rho g \pi r^2 h = 0 \Rightarrow h = \frac{2\gamma \cos\theta}{\rho g r} \tag{1.210}$$

If the liquid does not wet the solid surface and has a high surface tension like that of mercury, then the contact angle is greater than $\pi/2$ ($\cos\theta < 0$), the liquid is depressed inside the capillary tube, and the height of liquid inside the capillary tube is negative ($h < 0$) (Fig. 1.33b). Note that the terms of (1.209) keep the same signs as when there is capillary rise.

1.17 Membrane Potential

Consider a closed system at constant temperature and volume, consisting of parts 1 and 2, with potassium chloride aqueous solutions in both of these parts and the potassium chloride molality greater in part 1. Assume that the membrane that separates parts 1 and 2 is permeable to potassium ions but prevents chloride ions from passing through (Fig. 1.34). The transport of potassium ions through a porous membrane from higher concentration (part 1) to lower concentration (part 2) leaves an excess of chloride ions in part 1 and an excess of potassium ions in part 2. As a consequence of this, the solutions of parts 1 and 2 become no longer electrically neutral, and an electric potential is formed that opposes the continued migration of potassium ions from part 1 to part 2 (Fig. 1.34). Note that the electrostatic interactions between ions of a solution are very strong. Small variations in the charge density lead to strong electrostatic forces that tend to restore electric neutrality. This means that the bulk of solutions 1 and 2 are, to a good approximation, electrically neutral, and the electrical potential difference due to potassium ion transport through the membrane is mostly located in the membrane interfacial surfaces liquid–membrane and membrane–liquid (Fig. 1.34). The electric potential due to the excess of charge on both sides of the membrane eventually prevents the

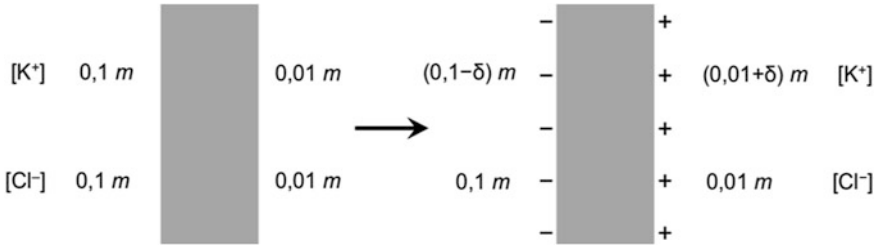


Fig. 1.34 Irreversible transport of potassium ions through a porous membrane destroys electrical neutrality in the liquid–membrane and membrane–liquid interfaces

continuation of charge transport through the membrane, leading the process of transport of potassium ions to an end.

In order to interpret the irreversible process of potassium ion transport through the porous membrane, we start by considering parts 1 and 2 with amounts of potassium ion n_1 and n_2, chemical potentials μ_1 and μ_2, and electric potentials ϕ_1 and ϕ_2. Since we have a closed system, the total amount of potassium ions is fixed and we can write

$$-\mathrm{d}n_1 = \mathrm{d}n_2 \equiv \mathrm{d}\xi \tag{1.211}$$

where ξ represents the extent of potassium ions passing through the membrane. For each $\mathrm{d}n_i$, besides the chemical potential associated with mass transfer, there is an additional contribution of electric energy (= charge $\times$ electric potential) associated with the charge transfer. Representing by z the amount of transported ion charge and by F the Faraday constant, equal to the charge of 1 mole of elementary charges ($F = eN_\mathrm{A} = 96485.3365\ \mathrm{C\ mol^{-1}}$, see Appendix), we can write, at constant volume and temperature ($\mathrm{d}V = 0$, $\mathrm{d}T = 0$),

$$\mathrm{d}U = T\mathrm{d}S + (\mu_1 + zF\phi_1)\mathrm{d}n_1 + (\mu_2 + zF\phi_2)\mathrm{d}n_2 \tag{1.212}$$

Making use of the first equality of (1.115), Eq. (1.212) can be split into two equations, one for the exchange of energy *across system boundaries* (the whole system is kept at constant temperature),

$$\mathrm{d}U = T\mathrm{d}_\mathrm{e}S \tag{1.213}$$

the other being due to irreversible changes *inside* the system,

$$0 = T\mathrm{d}_\mathrm{i}S + (\mu_1 + zF\phi_1)\mathrm{d}n_1 + (\mu_2 + zF\phi_2)\mathrm{d}n_2 \tag{1.214}$$

Note that addition of (1.213) and (1.214) gives (1.212). Substitution of (1.211) in (1.214) and rearrangement of the resulting equation leads to

$$T\mathrm{d}_\mathrm{i}S = (\mu_1 + zF\phi_1)\mathrm{d}\xi - (\mu_2 + zF\phi_2)\mathrm{d}\xi \tag{1.215}$$

This equation can be written as

$$T\mathrm{d}_\mathrm{i}S = (\tilde{\mu}_1 - \tilde{\mu}_2)\mathrm{d}\xi = \widetilde{A}\mathrm{d}\xi \tag{1.216}$$

where $\tilde{\mu}$ is the **electrochemical potential**, defined by

$$\widetilde{\mu}_i = \mu_i + zF\phi_i \tag{1.217}$$

and $\tilde{A}$ represents the **electrochemical affinity**, equal to the difference of electrochemical potentials of parts 1 and 2,

$$\tilde{A} = \tilde{\mu}_1 - \tilde{\mu}_2 = (\mu_1 - \mu_2) + zF(\phi_1 - \phi_2) \tag{1.218}$$

(see *Kondepudi*, Further Reading). The transport of potassium ions through the membrane is an irreversible process. Hence, $d_iS > 0$ and from (1.216) we can write

$$\frac{d_iS}{dt} = \frac{(\tilde{\mu}_1 - \tilde{\mu}_2)}{T}\frac{d\xi}{dt} = \frac{\tilde{A}}{T}\frac{d\xi}{dt} \tag{1.219}$$

where dt is the time interval that corresponds to the entropy production d_iS.

At equilibrium, $d_iS/dt = 0$, and both the electrochemical affinity $\tilde{A}$ and the rate of the irreversible process $d\xi/dt$ are zero ($d\xi/dt = 0$ is a consequence of $\tilde{A}$ being zero). Then, at equilibrium, the electrochemical potentials are equal,

$$\tilde{\mu}_1 - \tilde{\mu}_2 = 0 \tag{1.220}$$

In turn, (1.218) and (1.220) lead to

$$\phi_2 - \phi_1 = \frac{1}{zF}(\mu_1 - \mu_2) \tag{1.221}$$

where the difference between electrical potentials is called the **membrane potential** (a positive quantity). Substitution of (1.160) in (1.221) leads to

$$\phi_2 - \phi_1 = \frac{RT}{zF}\ln\left(\frac{a_1}{a_2}\right) \tag{1.222}$$

A porous membrane that allows migration of a specific ion is called an **ion-selective membrane**. The mechanism of selective migration can be realized with a ligand embedded in the membrane, which selectively coordinates the specific ion, thus promoting its transport through the membrane (Fig. 1.35).

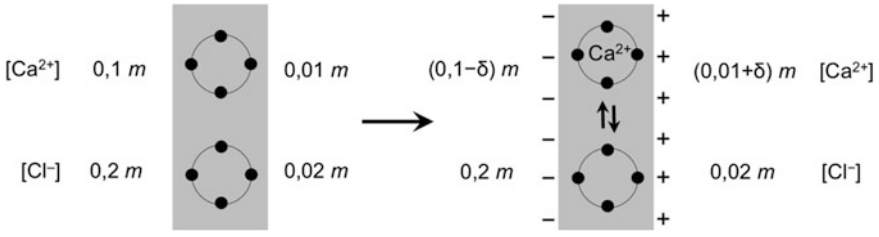


Fig. 1.35 The selective irreversible transport of calcium ions through the membrane is a result of their coordination by a ligand embedded in the membrane

1.18 Electrochemical Cell

Just try to submerge a copper wire into an aqueous solution of zinc sulfate (colorless). Nothing happens! However, if a zinc plate is immersed in a copper sulfate solution (blue), the solution gradually becomes colorless, the zinc plate becomes progressively eroded, and copper is deposited over it. This experiment can be described by the following chemical equation,

$$\mathrm{Zn(s)} + \mathrm{Cu^{2+}(aq)} \rightarrow \mathrm{Zn^{2+}(aq)} + \mathrm{Cu(s)} \tag{1.223}$$

which can be decomposed into the **oxidation half-reaction**

$$\mathrm{Zn(s)} \rightarrow \mathrm{Zn^{2+}(aq)} + 2e^- \tag{1.224}$$

and the **reduction half-reaction**

$$\mathrm{Cu^{2+}(aq)} + 2e^- \rightarrow \mathrm{Cu(s)} \tag{1.225}$$

The electrons produced in the oxidation half-reaction are consumed in the reduction half-reaction. These half-reactions can be carried out in separate beakers (zinc plate and zinc sulfate solution in one beaker, copper plate and copper sulfate solution in another) with a **salt bridge** connecting the solutions and an external circuit connecting the **electrodes** (from *electric* + Greek *hodos* *"way"*), as schematically shown in Fig. 1.36. The electrons produced by the oxidation half-reaction in the zinc plate go through the **external circuit** toward the copper plate, where they are combined with copper ions from the solution and deposited over the copper electrode (copper ion reduction half-reaction). The whole apparatus, an **electrochemical cell** that converts the Gibbs energy from the oxidation–reduction reaction (1.223) into electric energy, is called a **galvanic cell** (from Luigi Galvani, 1737–1798). The electric energy from a galvanic cell can be used to feed an electric motor in an external circuit. An electrochemical cell that performs the

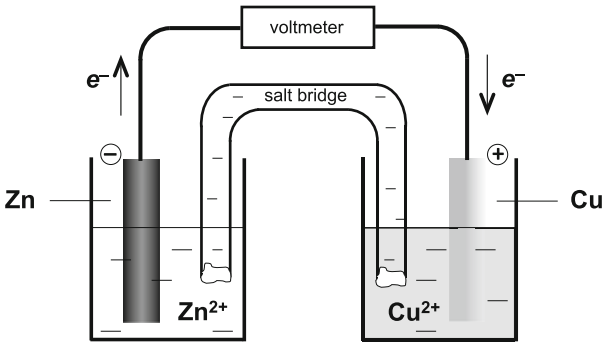


Fig. 1.36 Galvanic cell Zn(s) | Zn^{2+}(aq) || Cu^{2+}(aq) | Cu (schematic)

opposite transformation, that is, uses electric energy to drive an oxidation–reduction reaction, is called an **electrolytic cell**. Galvanic and electrolytic cells are electrochemical cells. From now on, we consider only galvanic cells.

The electrode where the oxidation occurs (**anode**, from Greek *ana* + *hodos* = *"up"* + *"way"*) becomes negative due to the electrons produced in the oxidation half-reaction. In turn, the electrode where the reduction occurs (**cathode**, from Greek *kata* + *hodos* = *"down"* + *"way"*) consumes electrons, thus becoming positive. In the external circuit, electrons migrate from the anode to the cathode. Thus, positive current flows to the anode as suggested by the origin of this word.

A salt bridge provides electrical contact between the oxidation and the reduction half-cells of a galvanic cell. The ideal salt bridge electrolyte is chosen so that the electric potential differences across both salt bridge–solution junctions cancel each other, so that they do not affect the electrochemical potential difference. For a 1:1 salt bridge electrolyte, both ions should have approximately the same migratory speed.

An electrochemical cell can be represented in a convenient and quick way by a **cell diagram**, where one vertical bar | represents an interface, and two vertical bars ‖ stand for the salt bridge junctions. The cathode half-cell is on the right of the cell diagram. The following cell diagram,

$$\mathrm{Zn(s)}\left|\mathrm{Zn^{2+}(aq)}\right|\left|\mathrm{Cu^{2+}(aq)}\right|\mathrm{Cu(s)} \tag{1.226}$$

represents the galvanic cell of Fig. 1.36.

1.18.1 Nernst Equation

The irreversible processes of an ion-selective membrane and a galvanic cell create an electric potential difference. Just as for an ion-selective membrane, the rate of entropy production of a galvanic cell is given by

$$\frac{\mathrm{d_i}S}{\mathrm{d}t} = \frac{\widetilde{A}}{T}\frac{\mathrm{d}\xi}{\mathrm{d}t} \tag{1.227}$$

where $\tilde{A}$ is the electrochemical affinity of the galvanic cell,

$$\frac{\mathrm{d}\xi}{\mathrm{d}t} = \frac{I}{nF} \tag{1.228}$$

I is the electric current intensity, that is, the charge transported per unit of time, and n is the number of electrons involved in the corresponding oxidation–reduction reaction. In dealing with a galvanic cell, we use a procedure similar to the one

followed for an ion-selective membrane and begin by evaluating the electrochemical affinity (see *Kondepudi*, Further Reading).

Let us consider a galvanic cell in which the following oxidation and reduction half-reactions take place:

$$\mathrm{X} \rightarrow \mathrm{X}^{n+} + ne^- \quad \mathrm{Y}^{n+} + ne^- \rightarrow \mathrm{Y} \tag{1.229}$$

For each of these half-reactions, the electrochemical affinity is the difference between the electrochemical potentials of the "reactants" (chemical species in the first member) and the electrochemical potentials of the "products" (chemical species in the second member). The chemical species X, X^{n+}, Y^{n+}, and Y are assigned to the corresponding chemical potentials. In turn, in the external circuit of the galvanic cell, the chemical potential of the electron μ_e is a function of the electronic density and temperature, physical quantities we assume to be constant. Hence, μ_e is constant, and the electrochemical potential for each electron is given by

$$\widetilde{\mu}_\mathrm{e} = \mu_\mathrm{e} - F\phi \tag{1.230}$$

where ϕ is the electric potential, and the minus sign is due to the negative charge of the electron. The electrochemical affinities of the anode and cathode half-reactions are given by

$$\begin{array}{ll} \mathrm{X} \rightarrow \mathrm{X}^{n+} + ne^- & \widetilde{A}^a = \mu_\mathrm{X} - [\mu_{\mathrm{X}^{n+}} + n(\mu_\mathrm{e} - F\phi^a)] \\ \mathrm{Y}^{n+} + ne^- \rightarrow \mathrm{Y} & \widetilde{A}^c = [\mu_{\mathrm{Y}^{n+}} + n(\mu_\mathrm{e} - F\phi^c)] - \mu_\mathrm{Y} \end{array} \tag{1.231}$$

where the superscripts a and c stand for anode and cathode, respectively. Since the oxidation–reduction reaction is the sum of the oxidation and reduction half-reactions, the electrochemical affinity of a galvanic cell $\tilde{A}$ is also the sum of the electrochemical affinities of the anode and cathode reactions,

$$\begin{array}{l} \mathrm{X} + \mathrm{Y}^{n+} \rightarrow \mathrm{X}^{n+} + \mathrm{Y} \\ \widetilde{A} = \widetilde{A}^a + \widetilde{A}^c = (\mu_\mathrm{X} + \mu_{\mathrm{Y}^{n+}}) - (\mu_{\mathrm{X}^{n+}} + \mu_\mathrm{Y}) - nF(\phi^c - \phi^a) \end{array} \tag{1.232}$$

This equation should be compared with (1.218), which refers to the electrochemical affinity of an ion-selective membrane. In the latter, the irreversible process is a physical process consisting in the transfer of an ion through the semipermeable membrane, whereas in the electrochemical cell, the irreversible process is an oxidation–reduction reaction of a galvanic cell.

When $\tilde{A} = 0$, then $\mathrm{d}\xi/\mathrm{d}t = 0$ [see (1.227)], and the current intensity is zero [see (1.228)]. Hence,

$$\Delta\varepsilon = \phi^\mathrm{c} - \phi^\mathrm{a} = \frac{1}{nF}[(\mu_\mathrm{X} + \mu_{\mathrm{Y}^{n+}}) - (\mu_{\mathrm{X}^{n+}} + \mu_\mathrm{Y})] \tag{1.233}$$

where $\Delta\varepsilon = \phi^{c} - \phi^{a}$, the potential difference between cathode and anode at zero-current intensity, is called the **zero-current cell potential** [compare (1.233) with (1.221) for the membrane potential]. Note that every electrochemical cell has an internal electric resistance R_{int}. If the electrical current across the cell is I, then the potential IR_{int} opposes the electrical potential of the cell. Hence, the zero-current situation ensures that its internal resistance does not affect the cell potential, since the product IR_{int} becomes zero when $I = 0$.

As is well known, chemical potentials depend on the nature and activities of chemical species according to the following expression:

$$\mu_k = \mu_k^0 + RT \ln a_k \tag{1.234}$$

Substitution of this expression in (1.233), with k taking the values X, Y^{n+}, X^{n+}, and Y, leads to the **Nernst equation**

$$\Delta\varepsilon = \Delta\varepsilon^0 - \frac{RT}{nF} \ln\left(\frac{a_{X^{n+}} a_Y}{a_X a_{Y^{n+}}}\right) \tag{1.235}$$

where

$$\Delta\varepsilon^0 = \frac{1}{nF}\left[\left(\mu_X^0 + \mu_{Y^{n+}}^0\right) - \left(\mu_{X^{n+}}^0 + \mu_Y^0\right)\right] \tag{1.236}$$

is the **standard cell potential**. When chemical equilibrium is reached, the reaction quotient, that is, the expression within parentheses in (1.235), becomes equal to the equilibrium constant, and $\Delta\varepsilon = 0$. Then we can write

$$\Delta\varepsilon^0 = \frac{RT}{nF} \ln K_e \tag{1.237}$$

At 298.15 K,

$$\log K_e \approx 16.9 n \Delta\varepsilon^0 \tag{1.238}$$

For $n = 1$ and $\Delta\varepsilon^0 = 0.5$ V, $\log K_e \approx 8.5$, that is, $K_e \approx 10^{8.5} \approx 3 \times 10^8$.

If the Nernst equation (1.235) is rewritten so that the cathode and anode potentials are separate terms, one obtains

$$\Delta\varepsilon = \left(\varepsilon^{0,c} - \frac{RT}{nF} \ln \frac{a_Y}{a_{Y^{n+}}}\right) - \left(\varepsilon^{0,a} - \frac{RT}{nF} \ln \frac{a_X}{a_{X^{n+}}}\right) \tag{1.239}$$

where the expressions in parentheses refer to reduction half-reactions for the cathode and anode. Note that both arguments of the logarithm functions correspond to half-reactions written as reductions. Therefore, we can generalize and write

$$X^{n+} + ne^- \rightarrow X \quad \varepsilon = \varepsilon^0 - \frac{RT}{nF} \ln \frac{a_X}{a_{X^{n+}}} \tag{1.240}$$

where ε is the **electrode potential** and ε^0 is the **standard electrode potential**, which is the electrode potential for unit activities of X^{n+} and X (**E19**, **E20**).

Notes

§1. Pressure is defined as force per unit area. The SI unit of pressure is the pascal, 1 Pa = 1 N m^{-2}. A column of liquid of uniform density ρ and height h exerts a pressure ρhg, where g represents the acceleration due to gravity; 1 mmHg is the pressure exerted by a column of mercury of height 1 mm when g = 9.80665 m s^{-2} and ρ(Hg) = 13.5951 g cm^{-3}. 1 atm = 760 mmHg = 101325 Pa, 1 bar = 100,000 Pa. Temperature is usually expressed in kelvin (K), degrees Celsius (°C), or degrees Fahrenheit (°F). The Celsius and Fahrenheit scales are empirical, and use the melting and boiling temperatures of water at 1 atm of external pressure as references. For the Celsius scale, 0 °C corresponds to the normal melting point of water, and 100 °C corresponds to the normal boiling point of water. In the Fahrenheit scale, 32 °F corresponds to the normal melting point of water, and 212 °F corresponds to the normal boiling point of water. The Kelvin scale is defined using the second law of thermodynamics, namely, the efficiency of a reversible heat engine or Carnot engine. The triple point of water is exactly 273.16 K (0.01 °C) and 611.73 Pa. The Kelvin and Celsius scales are related by T/K = T/°C + 273.15.

§2. Let $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ be the velocities of the test molecule O_2 and of one target molecule N_2. The relative velocity is equal to $\boldsymbol{v}_1 - \boldsymbol{v}_2$. The absolute value of this vector is given by

$$\sqrt{(\boldsymbol{v}_1 - \boldsymbol{v}_2) \cdot (\boldsymbol{v}_1 - \boldsymbol{v}_2)} = \sqrt{v_1^2 + v_2^2 - 2\boldsymbol{v}_1 \cdot \boldsymbol{v}_2}$$

Taking ensemble averages, we can write

$$\sqrt{\langle (\boldsymbol{v}_1 - \boldsymbol{v}_2) \cdot (\boldsymbol{v}_1 - \boldsymbol{v}_2) \rangle} = \sqrt{\langle v_1^2 \rangle + \langle v_2^2 \rangle - 2\langle \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 \rangle}$$

The molecular motions are random, and so the velocities of molecules 1 and 2 are uncorrelated, that is,

$$\langle \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 \rangle = 0$$

In addition, the space of molecular velocities is isotropic,

$$\langle v_1^2 \rangle = \langle v_2^2 \rangle = \langle v^2 \rangle$$

Hence,

$$\sqrt{\langle (\boldsymbol{v}_1 - \boldsymbol{v}_2) \cdot (\boldsymbol{v}_1 - \boldsymbol{v}_2) \rangle} = \sqrt{\langle v_1^2 \rangle + \langle v_2^2 \rangle - 2\langle \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 \rangle} = \sqrt{2}\sqrt{\langle v^2 \rangle} = \sqrt{2} v_{rms}$$

§3. In an adiabatic transformation (from Greek, *adiábatos*, *"impassable"*), the system is thermally isolated from the surroundings, that is, there is no heat exchanged between the system and its surroundings. The ideal gas equation for an adiabatic transformation is

$$TV^{\gamma - 1} = \text{constant}$$

Substitution of T by $pV/(nR)$ leads to the equation

$$pV^{\gamma} = \text{constant}$$

where γ is defined by

$$\gamma = C_{\mathrm{m}p}/C_{\mathrm{m}V}$$

and $C_{\mathrm{m}p}$ and $C_{\mathrm{m}V}$ are the isobaric and isochoric molar heat capacities. For a monatomic ideal gas, $\gamma = 5/3$.

§4. In the following derivation, we use the definition of mechanical work given by (1.84) and the expression for the first law that results from integration of both members of (1.187), $\Delta U = Q + W$. We assume 1 mole of ideal gas, so we deal with molar volumes. A Carnot cycle consists of an ideal gas undergoing the following four steps:

Step A→B (Isothermal Expansion)

In this step, the gas absorbs Q_{AB} from the heat source at temperature θ_{h} and uses this heat to do the work W_{AB} on the surroundings. Note that the energy of an ideal gas depends only on its temperature [see (1.19)]. Hence, for an isothermal transformation, the energy change is zero, and so from the first law, $Q_{AB} + W_{AB} = 0$:

$$W_{AB} = -\int_A^B p\mathrm{d}V = -R\theta_{\mathrm{h}}\int_A^B \frac{1}{V}\mathrm{d}V = -R\theta_{\mathrm{h}}\ln(V_B/V_A)$$

For an expansion, $V_B > V_A$, and so W_{AB} is a negative quantity (work done by the gas on the surroundings).

Step B → C (Adiabatic Expansion)

Since this step is adiabatic, $Q_{BC} = 0$ and

$$p_B V_B^{\gamma} = p_C V_C^{\gamma}$$

Using this equality, the work done by the system is given by

$$W_{BC} = -\int_B^C p\mathrm{d}V = -\int_B^C \frac{pV^{\gamma}}{V^{\gamma}}\mathrm{d}V = -\frac{p_C V_C^{\gamma} V_C^{1-\gamma}}{1-\gamma} + \frac{p_B V_B^{\gamma} V_B^{1-\gamma}}{1-\gamma} = -\frac{p_C V_C}{1-\gamma} + \frac{p_B V_B}{1-\gamma}$$
$$= -\frac{R(\theta_{\mathrm{c}} - \theta_{\mathrm{h}})}{1-\gamma} = -\frac{R(\theta_{\mathrm{h}} - \theta_{\mathrm{c}})}{\gamma - 1}$$

where W_{BC} is a negative quantity (this step corresponds to an expansion, $\theta_B = \theta_{\mathrm{h}} > \theta_{\mathrm{c}} = \theta_{\mathrm{C}}$ and $\gamma > 1$).

Step C → D (Isothermal Compression)

In this step, the work W_{CD} done on the system during the gas compression is given by

$$W_{CD} = -\int_C^D p\mathrm{d}V = -R\theta_{\mathrm{c}}\int_C^D \frac{1}{V}\mathrm{d}V = -R\theta_{\mathrm{c}}\ln(V_D/V_C)$$

Being a compression, $V_D < V_C$, and W_{CD} is a positive quantity. For an isothermal transformation in an ideal gas, the energy change is zero, and so $Q_{CD} + W_{CD} = 0$ (Q_{CD} corresponds to released heat, and is thus a negative quantity).

Step D → A (Adiabatic Compression)

Since this step is adiabatic, $Q_{DA} = 0$ and

$$p_D V_D^{\gamma} = p_A V_A^{\gamma}$$

Using this equality, the work done on the system is given by

$$\begin{aligned} W_{DA} &= -\int_D^A p\mathrm{d}V = -\int_D^A \frac{pV^{\gamma}}{V^{\gamma}}\mathrm{d}V = -\frac{p_A V_A^{\gamma} V_A^{1-\gamma}}{1-\gamma} + \frac{p_D V_D^{\gamma} V_D^{1-\gamma}}{1-\gamma} \\ &= -\frac{p_A V_A}{1-\gamma} + \frac{p_D V_D}{1-\gamma} = \frac{R(\theta_{\mathrm{h}} - \theta_{\mathrm{c}})}{\gamma - 1} \end{aligned}$$

From the results obtained for W_{BC} and W_{DA}, we conclude that $W_{BC} + W_{DA} = 0$.

The work done on the Carnot reversible cycle is equal to the sum of the works done in the four steps of the cycle. Using the above results ($Q_{AB} + W_{AB} = 0$, $Q_{BC} = 0$, $Q_{CD} + W_{CD} = 0$, $Q_{DA} = 0$, and $W_{BC} + W_{DA} = 0$), we can write

$$W_{AB} + W_{BC} + W_{CD} + W_{DA} = W_{AB} + W_{CD} = -Q_{AB} - Q_{CD}$$

Hence, the efficiency for the Carnot cycle is given by

$$\begin{aligned} \eta_{\mathrm{Carnot}} &= 1 - \frac{Q_{\mathrm{c}}}{Q_{\mathrm{h}}} = \frac{W}{Q_{\mathrm{h}}} = \frac{Q_{AB} - |Q_{CD}|}{Q_{AB}} = 1 - \frac{|Q_{CD}|}{Q_{AB}} = 1 + \frac{W_{CD}}{W_{AB}} \\ &= 1 + \frac{R\theta_{\mathrm{c}} \ln(V_D/V_C)}{R\theta_{\mathrm{h}} \ln(V_B/V_A)} = 1 - \frac{R\theta_{\mathrm{c}} \ln(V_C/V_D)}{R\theta_{\mathrm{h}} \ln(V_B/V_A)} = 1 - \frac{\theta_{\mathrm{c}}}{\theta_{\mathrm{h}}} \end{aligned}$$

where W, Q_{h}, and Q_{c} represent absolute values. The equations for the isothermal transformations ($p_{\mathrm{A}}V_{\mathrm{A}} = p_{\mathrm{B}}V_{\mathrm{B}}$ and $p_{\mathrm{C}}V_{\mathrm{C}} = p_{\mathrm{D}}V_{\mathrm{D}}$) and adiabatic transformations ($p_{\mathrm{B}}V_{\mathrm{B}}^{\gamma} = p_{\mathrm{C}}V_{\mathrm{C}}^{\gamma}$ and $p_{\mathrm{D}}V_{\mathrm{D}}^{\gamma} = p_{\mathrm{A}}V_{\mathrm{A}}^{\gamma}$) can be simultaneously solved, provided $V_{\mathrm{B}}/V_{\mathrm{A}} = V_{\mathrm{C}}/V_{\mathrm{D}}$. This equality justifies the last equation above. Therefore, the Carnot engine efficiency is given by

$$\eta_{\mathrm{Carnot}} = 1 - \frac{Q_{\mathrm{c}}}{Q_{\mathrm{h}}} = 1 - \frac{\theta_{\mathrm{c}}}{\theta_{\mathrm{h}}}$$

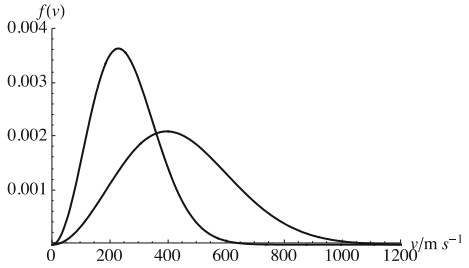
Mathematica Codes

M1. Maxwell Distribution of Molecular Speeds

```
Clear[avogadro,kB,m]
norm=1/Integrate[Exp[-m*v^2/(2*kB*T)]*4*Pi*v^2,{v,0,Infinity},Assumptions→{T>0,m>0,kB>0}]
avogadro=6.022 10^23;(* molecules mol^-1 *)
kB=1.381 10^-23;      (* J K^-1 *)
m=0.032/avogadro;     (* kg molecule^-1 *)
norm
Plot[norm*Exp[-m*v^2/(2*kB*T)]*4*Pi*v^2/.T→{100,300},{v,0,1200},
    PlotRange→{{0,1200},{0,0.004}},AxesLabel→{"v/m s^-1","f(v)"},
    AxesStyle→Directive[FontSize→14],PlotStyle→{Black,{Thickness[0.006]}}]
```

$$\frac{1}{2\sqrt{2}\,\pi^{3/2}\left(\frac{kB\,T}{m}\right)^{3/2}}$$

$$\frac{0.0000151549}{T^{3/2}}$$



This *Mathematica* code evaluates the normalization factor of `f(v)` in (1.27) and plots this probability function for a gas of oxygen molecules at 100 and 300 K. The normalization factor is calculated as the inverse of the definite integral of `f(v)` in the range `{0,∞}` and requires the use of the *Mathematica* command `Integrate[f(v),{v,0,Infinity},Assumptions→{T>0,m>0,kB>0}]` where `v` is the integration variable, whose range is between zero and infinity, because `v` represents a molecular speed, that is, an absolute value of the velocity vector, and the option `Assumptions` indicates that the symbolic quantities `T`, `m`, and `kB` are positive. They are assigned specific values in lines of code 3, 4, and 5, and line of code 6 outputs the value of the `norm` for the oxygen molecule, at an unspecified temperature `T`.

The *Mathematica* command `Plot[f[v]/.T →{100,300},{v,0,1200}]` generates plots of functions `f(v)` in the range `{0,1200}`, where `T` is replaced by 100 and 300 (/. stands for the `ReplaceAll` *Mathematica* command). This is the point where the variable `T` is assigned specific values, namely, 100 and 300. The physical quantities `m`, `v`, and `T`, and the constants `avogadro` and `kB` are in SI

units. If you wish to run the above code for different molecular masses or different temperatures and do not want to transfer universal constants and mass values from one run to the next, you should use `Clear[avogadro,kB,m]` as the first line of the code. This Mathematica function clears values and definitions of the specified symbols, making it possible to evaluate the `norm` on the second line of code using symbolic constants and variables.

Suggestion: Use the `Graphics` primitive `Text` to add the temperatures "100 K" and "300 K" next to the corresponding curves in the above plot. If necessary, consult Wolfram Documentation in the Help menu of *Mathematica* with the words `Graphics` and `Text`.

M2. Critical Point for Van der Waals Fluid

```
p[V_,T_]:=R T/(V-b)-a/V^2;
sol=Solve[D[p[V,T],V]==0&&D[p[V,T],{V,2}]==0,{V,T}]
sol[[1]]
Print["Critical Volume"]
V/.sol[[1]]
Print["Critical Temperature"]
T/.sol[[1]]
Print["Critical Pressure"]
p[V/.sol[[1]],T/.sol[[1]]]
Print["van der Waals Equation in Reduced Variables"]
Solve[3b==1&&8a/(27bR)==1&&a/(27b^2)==1,{a,b,R}][[1]]
```

$$\left\{\left\{V \to 3\,b,\ T \to \frac{8\,a}{27\,b\,R}\right\}\right\}$$

$$\left\{V \to 3\,b,\ T \to \frac{8\,a}{27\,b\,R}\right\}$$

```
Critical Volume
```

$$3\,b$$

```
Critical Temperature
```

$$\frac{8\,a}{27\,b\,R}$$

```
Critical Pressure
```

$$\frac{a}{27\,b^2}$$

```
van der Waals Equation in Reduced Variables
```

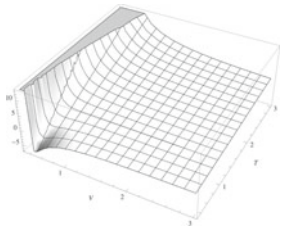
$$\left\{a \to 3,\ b \to \frac{1}{3},\ R \to \frac{8}{3}\right\}$$

This code shows how to use *Mathematica* to determine the critical values of volume, temperature, and pressure for a Van der Waals fluid. The code solves the system of Eqs. (1.49) and confirms the critical values presented in (1.50). The first line of code defines the `p[V,T]` function for the Van der Waals equation, and the

second line of code solves the system of Eqs. (1.49) to determine the critical values of volume and temperature. The expression `D[p[V,T],V]` gives the partial derivative of `p` with respect to `V` at constant `T`, and `D[p[V,T],{V,2}]` gives the second-order partial derivative of `p` with respect to `V` at constant `T`. The solutions of the system of equations are assigned to `sol`, which is a list of a list of length 2 (see the first line of output; in *Mathematica*, a list is denoted by {…}). The second line of output shows that `sol[[1]]` stands for the first part of `sol`, that is, a list of the critical volume and critical temperature. The `Print` command of the fourth line of code is used to print the string of characters inside quotation marks (“`Critical Volume`”), and the next line of code uses `ReplaceAll (/.)` to replace `V` by the first part of `sol`, that is, by `sol[[1]]`. The last line of code solves for the values of `a`, `b`, and `R` that lead to critical values of volume, temperature, and pressure equal to 1, thus to the Van der Waals equation in reduced variables.

M3. 3D Plot of the Van der Waals Equation in Reduced Variables

```
Plot3D[8T/3/(V-1/3)-3/V^2,{V,0.34,3},{T,0.34,3},AxesLabel→{V,T},
PlotStyle→Gray,Lighting→"Neutral"]
```



This *Mathematica* code plots the Van der Waals equation in reduced variables [see (1.53)] using the *Mathematica* command

`Plot3D[f[V,T],{V,0.34,3},{T,0.34,3}]`

where `f[V,T]` corresponds to the reduced pressure as a function of the reduced volume `V` and reduced temperature `T`, in the above indicated ranges. Note that a function of two independent variables is represented by a surface. The Van der Waals isotherms of Fig. 1.11 are cross sections of this surface by planes that correspond to defined temperature values. The option `Lighting` specifies the simulated lighting used to color the plotted surface. `Lighting→“Neutral”` simulates white light.

Suggestion: Run the above *Mathematica* code line leaving out the options `PlotStyle` and `Lighting`.

M4. Absolute and Reduced Temperatures for the Van der Waals Carbon Dioxide

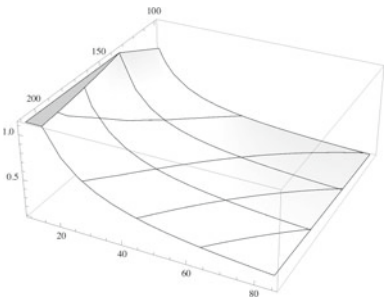
```
Print["Critical Temperature for Carbon Dioxide / K"]
Tc=SetPrecision[8*3.658/(27*0.0429*0.083145),4]
Print["Carbon Dioxide - List of {Absolute Temp,Reduced Temp}]
{{274,274/Tc},{284,284/Tc},{294,294/Tc},{304,304/Tc},{314,314/Tc},{324,324/Tc},
    {334,334/Tc},{344,344/Tc}}
```

```
Critical Temperature for Carbon Dioxide / K

303.9

Carbon Dioxide - List of {Absolute Temp,Reduced Temp}

{{274, 0.9017}, {284, 0.9346}, {294, 0.9675},
 {304, 1.000}, {314, 1.033}, {324, 1.066}, {334, 1.099}, {344, 1.132}}
```

This *Mathematica* code evaluates the critical temperature for carbon dioxide (see the values of `a` and `b` in the Table for Van der Waals constants) and calculates reduced temperatures corresponding to absolute temperatures from 274 to 344 K in steps of 10.

M5. Isothermal and Adiabatic Transformations on an Ideal Gas Surface

```
R = 0.08206; (* L atm mol^-1 K^-1 *)
f[T_,V_]:=R*T/V
Plot3D[f[T,V],{T,210,100},{V,10,85},Mesh→{2,4},MeshFunctions→{#1&,#1*#2^(2/3)&},
  MeshStyle→{Black,Black},ViewPoint→{Pi,Pi/2,2},PlotStyle→Gray,Lighting→"Neutral"]
```



This *Mathematica* code plots the ideal gas surface with meshes for two isothermal and four adiabatic transformations. The option for `Plot3D` that specifies the number of isothermal and adiabatic transformations is

`Mesh→{2,4}`

and the option for the corresponding functions is

`MeshFunctions→{#1&,#1*#2^(2/3)&}`

where `#1&` is the pure function for an isothermal transformation (`#1` is the first argument of `f[T,V]`) and `#1*#2^(2/3)` is the pure function for an adiabatic transformation (`#2` is the second argument of `f[T,V]`). The adiabatic transformation is given by $TV^{\gamma-1} = constant$, where γ is equal to 5/3 for a monatomic ideal gas (see **§3**).

Suggestion: Change the above *Mathematica* code so that it runs for 35 isothermal transformations and 16 adiabatic transformations. Choose a `ViewPoint` from `Top`. Compare the obtained plot with that of Fig. 1.25.

M6. Efficiency of the Carnot Heat Engine

```
aa=Table[10^i,{i,0,3}]
bb=SetPrecision[Table[(1.-10^i/1001)*100,{i,0,3}],3]
cc=Riffle[aa,bb]
Partition[cc,2]
```

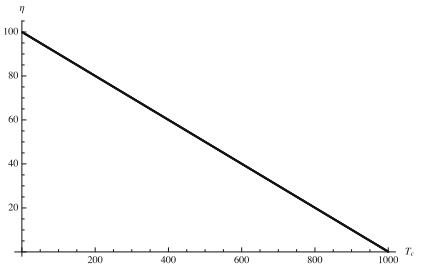
```
{1, 10, 100, 1000}
{99.9, 99.0, 90.0, 0.0999}
{1, 99.9, 10, 99.0, 100, 90.0, 1000, 0.0999}
{{1, 99.9}, {10, 99.0}, {100, 90.0}, {1000, 0.0999}}
```

This *Mathematica* code calculates the efficiencies of Carnot heat engines [see (1.107)] when the temperature of the hot reservoir is fixed at 1001 K and the temperatures of the cold reservoir are 1, 10, 100, and 1000 K. The efficiencies of Carnot engines are presented as percentages with three significant digits using the *Mathematica* command `SetPrecision`. This code illustrates the use of `Table` to obtain two lists of length 4: the list `aa` of four cold reservoir temperatures and the list `bb` composed of the corresponding Carnot efficiencies. These two lists are interleaved by the use of the *Mathematica* command `Riffle`. The command `Partition` partitions the list `cc` into nonoverlapping lists of length 2 composed of the temperature of the cold reservoir T_c and the maximum Carnot's engine efficiency η. Note that the above code can be compressed into the following single line of code:

```
Partition[Riffle[Table[10^i,{i,0,3}],
SetPrecision[Table[(1.-10^i/1001)*100,{i,0,3}],3]],2]
```

The following code uses `ListPlot` to plot the list of points (T_c, η), where T_c varies from 1 to 1000 in steps of 1, and η is indicated as a percentage. This plot shows a linear dependence of η versus T_c, with η decreasing with the increase of T_c.

```
ListPlot[Partition[Riffle[Table[i,{i,1,1000}],
    Table[(1.-i/1001)*100,{i,1,1000}]],2], PlotStyle→Black]
```



M7. Gibbs Energy and Affinity of a Chemical Reaction

Consider a hypothetical chemical reaction R(g) → P(g), where R and P are ideal gases and the equilibrium constant is assumed to be 3.0. Taking the Gibbs energy and the affinity as functions of the extent of the chemical reaction ξ [see (1.131) and (1.132)] whose range is [0,1], the partial derivative of the Gibbs energy with respect to ξ, at constant temperature and pressure, is equal to the negative of affinity of the chemical reaction [see second equality of (1.170)].

The following *Mathematica* code shows how the Gibbs energy of the above chemical reaction and the corresponding affinity evolve to equilibrium. In order to obtain the Gibbs energy as a function of ξ, we write

$$G = G_{\mathrm{R}}^{0} + \sum \Delta G$$

where G_{R}^{0} is the standard Gibbs energy of the reactant R (in the code, we assume $G_{\mathrm{R}}^{0} = 0$) and ΔG is obtained by substituting (1.188) into (1.185), that is, using

$$\Delta G = RT \ln\left(\frac{\chi}{K_{\mathrm{eq}}}\right)$$

where χ represents the reaction quotient

$$\chi = \frac{\xi}{1-\xi}$$

and K_{eq} is the equilibrium constant. In the code,

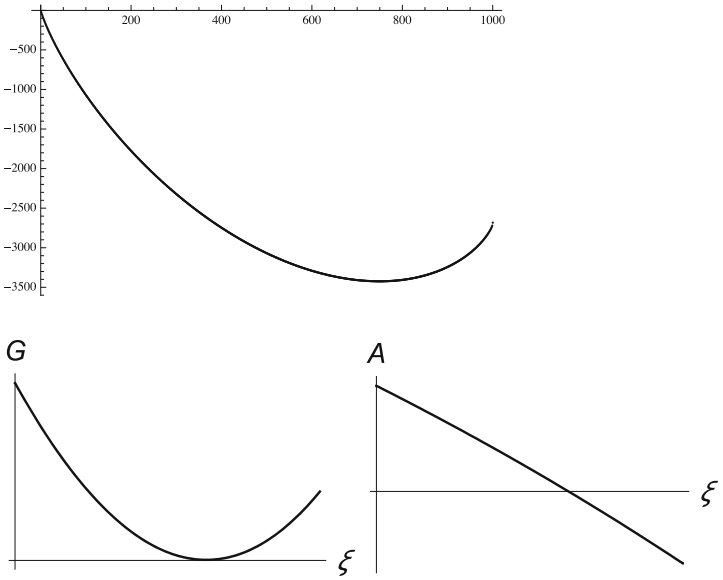
$$\xi = \frac{i}{1000}$$

where i varies in the range [1,1000]. Therefore, each step of ξ is equal to 0.001, and the command `ListPlot[data]` plots 1000 points. The *Mathematica* command `Fit` finds a least squares fit to `data` as a linear combination of functions in the list `{1,x,x^2,x^3,x^4}`, where `x` is real and corresponds to `i`. The value of `x` at equilibrium is obtained by the code lines

```
sol=Solve[bestfit'[x]==0,x][[3]];
eq=x/.sol;
```

Solving `bestfit'[x]==0` leads to three roots, the first two being complex numbers, the third being real and equal to 750.181. The third root, denoted by `sol`, is selected using `[[3]]` immediately after the `Solve` command. The value of `x` at equilibrium is assigned to `eq`. Plots of the Gibbs energy G and affinity A are restricted to the neighborhood of equilibrium.

```
(* R(g)→P(g) where R and P are ideal gases *)
R=0.008314; (* kJ mol^-1 K^-1 *)
T=298.15;   (* K *)
Keq=3; G=0;
data=Table[G=G+R*T*Log[i/1000./(1.-i/(1000.+0.0001))/Keq],{i,1,1000}];
(* extent of the chemical reaction ξ = i/1000. *)
(* 0.0001 prevents mathematical singularity at i = 1000 *)
(* argument of Log function is χ/Keq, where χ is the chemical reaction quotient *)
ListPlot[data,PlotStyle→Black]
Clear[bestfit]
bestfit[x_]=Fit[data,{1,x,x^2,x^3,x^4},x];
sol=Solve[bestfit'[x]==0,x][[3]];
eq=x/.sol;
(* Gibbsenergy and affinity plotted in the neighbourhood of equilibrium *)
Gibbsenergy=Plot[bestfit[x],{x,eq-50,eq+30},PlotStyle→{Black,Thickness[0.008]},Ticks→None,
    AxesLabel→{Style["ξ",Large,Black],Style["G",Large,Black]},ImageSize→{300,300}];
affinity=Plot[-bestfit'[x],{x,eq-50,eq+30},PlotStyle→{Black,Thickness[0.008]},Ticks→None,
    AxesLabel→{Style["ξ",Large,Black],Style["A",Large,Black]},ImageSize→{300,300}];
Row[{Gibbsenergy, affinity}, "   "]
```



Suggestion: Write a *Mathematica* code to obtain the Gibbs energy and the affinity as functions of the extent of the chemical reaction A + B → 2C and assume $K_{eq} = 100$.

Glossary

Activity Applies to chemical species in nonideal solutions and corrects mole fractions for departure from ideal behavior due to intermolecular interactions; it is obtained as the product of an activity coefficient by the mole fraction; see (1.161). Ideal behavior corresponds to the activity coefficient being equal to 1.

Adiabatic transformation A transformation in which the system is thermally isolated from the surroundings and so no heat is exchanged between the system and its surroundings; see **§3**.

Affinity of a chemical reaction The stoichiometric sum of chemical potentials of the reactants minus the stoichiometric sum of chemical potentials of the products; see (1.136). Affinity and Gibbs energy are related through the second equality of (1.170). At equilibrium, the Gibbs energy attains a minimum and the reaction affinity is zero; see Fig. 1.30 and *Mathematica* code **M7**.

Anode The electrode where oxidation occurs. In a galvanic cell, the Gibbs energy from an oxidation–reduction reaction is converted into electric energy. The oxidation at the anode produces electrons that give rise to a negative charged electrode and releases positive ions into solution, thus eroding the anode. In an electrolytic cell, the positive terminal of the applied voltage is directly connected to the anode, and so the anode becomes positive, causing oxidation to occur. The anode is where oxidation occurs, whether the cell is galvanic or electrolytic.

Carnot heat engine Consists of an ideal gas that operates over a reversible cycle formed by isothermal and adiabatic expansions followed by isothermal and adiabatic compressions; see Fig. 1.23 and *Mathematica* code **M5**. According to Carnot's theorem, the efficiency of the Carnot heat engine depends only on the temperatures of the hot and cold reservoirs; see (1.102).

Cathode The electrode where reduction occurs. In a galvanic cell, reduction at the cathode combines electrons from the electrode with positive ions from solution giving rise to neutral atoms that are deposited on the cathode and make it grow. In turn, depletion of electrons from the cathode makes it positive (the cathode has a higher electric potential than the anode). In an electrolytic cell, the negative terminal of the applied voltage is directly connected to the cathode, and so the cathode becomes negative, thus originating reduction. The cathode is where reduction occurs, whether the cell is galvanic or electrolytic.

Chemical potential The partial derivative of energy U with respect to the amount of substance n_k, at constant entropy S, constant volume V, and constant amounts of other system components $j \neq k$; see (1.139) and (1.140). Alternatively, the chemical potential is the partial derivative of the Gibbs energy G with respect to the amount of substance n_k, at constant temperature T, constant pressure p, and constant amounts of other system components $j \neq k$; see (1.169) and the first equation of (1.170). The chemical potential has units of energy per mole.

Closed system Can exchange energy, not matter, with its surroundings; see Fig. 1.15.

Electrochemical affinity For an ion-selective membrane is the difference of electrochemical potentials across the membrane; see (1.218) and Fig. 1.34. For an oxidation–reduction reaction, the electrochemical affinity is the stoichiometric sum of electrochemical potentials of the oxidation half-reaction plus the stoichiometric sum of electrochemical potentials of the reduction half-reaction; see (1.231) and (1.232). At equilibrium, the electrochemical affinity is zero: for an ion-selective membrane, this situation gives the membrane potential, see (1.221); for a galvanic cell, one obtains the zero-current cell potential; see (1.233).

Electrochemical cell A general expression that comprises galvanic and electrolytic cells.

Electrochemical potential The chemical potential (units of energy per mole) associated with mass transfer, plus electric energy per mole (= charge per mole x electric potential) associated with charge transfer; see (1.217).

Electrolytic cell Uses electric energy to drive an oxidation–reduction reaction.

Enthalpy A state function defined by $H \equiv U + pV$; see (1.92) and (1.93).

Entropy A state function usually represented by S that is central to the second law of thermodynamics. An arbitrary reversible cycle carried out in a given direction can be covered to the required precision by a mesh of infinitesimal Carnot cycles carried out

in the same direction. This enables one to conclude that the integral of dQ_{rev}/T over the arbitrary reversible cycle is zero, that is, $\mathrm{d}Q_{\mathrm{rev}}/T$ is the differential of a state function, called entropy; see (1.110) and (1.111). According to Ilya Prigogine (1917–2003, Nobel Prize in chemistry in 1977), a distinction should be made between two terms included in the entropy change $\mathrm{d}S$. The first, $\mathrm{d_e}S$, is the transfer of entropy across system boundaries due to energy and matter exchanges with the surroundings. This transfer of entropy can be performed reversibly, the corresponding entropy change d_eS being either positive or negative (no thermodynamic law can be formulated on this term). In turn, the second term, $\mathrm{d_i}S$, is the entropy produced within the system by irreversible processes that drive the system toward equilibrium and is always positive, $\mathrm{d_i}S > 0$, a statement that corresponds to the second law of thermodynamics.

Extent of a chemical reaction State variable whose differential is equal to the differential of the amount of a reactant or a product of a chemical reaction divided by the corresponding stoichiometric number; see (1.131) and (1.132). Stoichiometric numbers coincide with the stoichiometric coefficients for the products of a chemical reaction and are equal to the stoichiometric coefficients multiplied by −1 for the reactants. The definition of the extent of a chemical reaction results from the fact that infinitesimal arbitrary changes in the amounts of reactants and products are related to one another by the chemical reaction stoichiometry.

Galvanic cell Converts Gibbs energy from an oxidation–reduction reaction into electric energy.

Gibbs–Duhem equation Shows that the intensive variables temperature, pressure, and chemical potential are not independent variables; see (1.145). For one system component, see (1.146).

Gibbs energy State function defined by $G \equiv U + pV - TS$; see (1.166) and (1.168).

Gibbs phase rule Equates the number of degrees of freedom of a system at equilibrium with the number of independent components, minus the number of phases, plus 2; see (1.197).

Helmholtz energy State function defined by $F \equiv U - TS$; see (1.198) and (1.200).

Ideal gas Physical model of a gas whose equation of state is given by (1.5), (1.7), or (1.8). The particles of an ideal gas are assumed to be point masses (the total volume of the particles of an ideal gas is zero) and noninteracting.

Irreversible process Occurs in a specific direction, not in its reverse, and evolves to equilibrium. In the time interval dt during an irreversible process, $d_i S > 0$ and the rate of entropy production $d_i S/dt$ is positive; see (1.115) and (1.116).

Isobaric heat capacity Partial derivative of enthalpy H with respect to temperature T at constant pressure; see (1.95). The isobaric heat capacity has units of energy per kelvin.

Isochoric heat capacity Partial derivative of energy U with respect to temperature T at constant volume; see (1.94). The isochoric heat capacity has units of energy per kelvin.

Isolated system Does not exchange matter or energy with its surroundings; see Fig. 1.15.

Kinetic model of gases Describes a gas assuming that (i) the average distance between molecules is far greater than the molecules' dimensions, (ii) the molecules are in constant random motion, (iii) each molecule is seen as a sphere with elastic collisions with other molecules and the container walls, (iv) the molecules do not interact with each other except during collisions and (v) the gas molecules obey Newton's equations of motion. The distribution of molecular speeds is given by (1.27); see Fig. 1.6 for an illustration with oxygen molecules at 100 and 300 K.

Mean free path Average distance between gas molecule collisions. The mean free path is inversely proportional to the square of the cross section for collisions and to the gas number density; see (1.39).

Mechanical work In the time interval dt is defined by $\mathrm{d}W = -p\mathrm{d}V$; see (1.83) and Fig. 1.18.

Membrane potential Electric potential difference developed across a porous membrane, due to the transport of certain ions, from higher to lower electrochemical potentials; see (1.219).

Nernst equation Gives the zero-current cell potential as a function of the reaction quotient of an oxidation–reduction reaction expressed in terms of activities; see (1.235).

Number of degrees of freedom The number of independent intensive variables of a system at equilibrium.

Number of independent components For a system at equilibrium is the number of its components minus the number of distinct chemical equilibria involving them.

Open system Can exchange matter and energy with its surroundings; see Fig. 1.15.

Phases Distinct physical states of matter.

Surface tension Helmholtz energy per unit of interfacial area; see (1.205).

Van der Waals critical temperature Temperature at which the maximum and minimum of van der Waals isotherms coincide, giving rise to the inflection called the critical point and to the critical isotherm; see (1.49), and *Mathematica* code **M2**.

Van der Waals equation Equation of state for a fluid (gas and liquid) composed of particles with nonzero volumes and pairwise attractive interactions; see (1.48). This equation results from modification and improvement of the equation of state of an ideal gas. A table with Van der Waals constants of selected gases appears next to (1.48), and the Van der Waals isotherms for carbon dioxide can be seen in Fig. 1.9.

Van der Waals reduced variables Molar volume, temperature, and pressure of a Van der Waals fluid in units defined by the corresponding critical values; see (1.52). The Van der Waals equation in reduced variables is given by (1.53); see *Mathematica* codes **M3** and **M4.**

Exercises

E1. Use *Mathematica* to confirm the normalization factor for (1.23).
E2. Use *Mathematica* to confirm (1.29), (1.30), and (1.31) and evaluate the most probable speed, the mean molecular speed, and the root mean square speed, for O_2 at 298.15 K.
E3. Obtain estimates for the mean free path, the average distance between molecules, and the time for one collision of O_2 at 273.15 K and 1 atm of pressure. Assume that the radius of the collision cross section is equal to 0.3 nm.
E4. Use *Mathematica* to determine the critical values of volume, temperature, and pressure for the van der Waals equation.
E5. Use *Mathematica* to plot the isotherms of the van der Waals equation for carbon dioxide with interactive manipulation of the temperature value.
E6. Use *Mathematica* to generate a three-dimensional plot of the Van der Waals equation in reduced variables.
E7. Classify $4x^2y^3\ \mathrm{d}x + 3x^3y^2\ \mathrm{d}y$ as an exact or inexact differential.
E8. Show that $p\ \mathrm{d}V + V\ \mathrm{d}p$, with p and V being the pressure and volume variables, is an exact differential.
E9. Use the first law to arrive at

$$\mathrm{d}Q = \left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T + \left[p + \left(\frac{\partial U}{\partial V}\right)_T\right]\mathrm{d}V$$

and derive the following equality:

$$C_p - C_V = \left[p + \left(\frac{\partial U}{\partial V}\right)_T\right]\left(\frac{\partial V}{\partial T}\right)_p$$

E10. The heat capacity of iron is 25.09 J mol^{-1} K^{-1}. Calculate the entropy change for 4 mol of iron heated from 273 K to 373 K, assuming that the heat capacity for iron is constant in this temperature interval.
E11. Evaluate the standard reaction enthalpy for $C_3H_8(g) + 5O_2(g) \rightarrow 3CO_2(g) + 4H_2O(l)$, at 298.15 K $[\Delta H_f^o[C_3H_8(g)] = -104.7\,\mathrm{kJ\,mol^{-1}};\ \Delta H_f^o[CO_2(g)] = -393.5\,\mathrm{kJ\,mol^{-1}};\ \Delta H_f^o[H_2O(l)] = -285.8\,\mathrm{kJ\,mol^{-1}}]$.
E12. Rudolf Clausius expressed the second law in the following way: *No process is possible whose sole result is the transfer of heat from a body of lower temperature to a body of higher temperature*. In turn, Lord Kelvin used the following statement: *No process is possible in which the sole result is the absorption of heat from a reservoir and its complete conversion into work*. Show that Clausius's and Kelvin's statements correspond to equivalent heat engines.

E13. Consider an empirical, ideal gas thermometer defined by the volume as a function of the gas's Celsius temperature t at constant pressure, $V = V_0 + V_0 \, \alpha \, t$. How is V_0 defined? Determine α.
E14. Draw a Carnot cycle in the temperature–entropy space.
E15. Use *Mathematica* to determine the equilibrium constant for $2NO_2(g) \rightleftarrows N_2O_4(g)$ at 298.15 K $\{\Delta G_f^o[NO_2(g)] = 51.3\,kJ\,mol^{-1};\ \Delta G_f^o[N_2O_4(g)] = 99.8\,kJ\,mol^{-1}\}$ and the NO_2 and N_2O_4 mole fractions at equilibrium.
E16. Consider a one-component system with two phases in equilibrium. Apply the Gibbs–Duhem equation to each of the phases in equilibrium and determine the derivative of pressure with respect to temperature to obtain Clapeyron's equation. Apply this equation to the liquid–vapor transition and assume ideal behavior for the vapor phase to obtain the Clausius–Clapeyron equation.
E17. Apply the Gibbs phase rule to the equilibrium in a system containing iron, iron (III) oxide, carbon, oxygen, and carbon dioxide. Determine the number of degrees of freedom in the system and find the equations that involve the system's thermodynamic variables.
E18. Consider an irreversible process in an isolated system. Show that entropy tends to a maximum at equilibrium.
E19. Show that all the electrode potentials are defined for one electron involved in the corresponding reduction half-reaction.
E20. Evaluate the equilibrium constant for $2Ag^+ + Cu \rightarrow 2Ag + Cu^{2+}$ at 298.15 K.

References

Angrist SW (1968) Perpetual motion machines. Sci Am

Further Reading

Cheung C-K, Keough GE, Gross RH, Landraitis C (2009) Getting started with mathematica, third edn. Wiley, New Jersey
Kondepudi D (2008) Introduction to modern thermodynamics. Wiley, New Jersey
Kondepudi D, Prigogine I (1998) Modern thermodynamics: from heat engines to dissipative structures. Wiley, New Jersey
Wolfram S (2015) An elementary introduction to the Wolfram Language. Wolfram Media, Inc. Available online

Chemical Kinetics

2

Abstract

Chemical kinetics is the branch of chemistry that measures rates of chemical reactions, studies the factors that influence them, designs and prepares new catalysts, and interprets the results at the molecular level. The independent variable of chemical kinetics, from the chemical reaction starting moment when the reactants are mixed to its final moment when equilibrium is reached, is time, a variable introduced by the second law of thermodynamics for irreversible processes. The first study of the rate of a chemical reaction is credited to Ludwig Wilhelmy in 1850 for the decomposition of sucrose (table sugar) into glucose and fructose, in acid medium. Wilhelmy found that the rate of this chemical reaction is proportional to the existing amount of sucrose at each instant in the course of the chemical reaction. This chapter begins with sections on the rate of a chemical reaction, the experimental rate equation, and the effect of temperature change. We then consider elementary reactions, complex reactions, and extremely fast reactions. Most chemical reactions function like one-way streets: the concentrations of reactants decrease, those of reaction intermediates increase at first and decrease later, and the concentrations of products increase. However, for a few reactions far from equilibrium, the concentrations of some intermediate species oscillate, increasing and decreasing repeatedly. These reactions are illustrated with the Brusselator, a model chemical oscillator developed in the Brussels thermodynamic school founded by Prigogine. At the end of this chapter, the student can find two notes on matrix diagonalization and systems of first-order linear differential equations, useful for understanding the mathematical treatment given to the Brusselator, two Mathematica codes (First-Order Chemical Reaction, Brusselator) with references to expressions in the main text, detailed explanations for new commands and suggestions for the student to follow, a glossary that explains important scientific terms, and a list of exercises, whose complete answers can be found at the end of the book.

J.J.C. Teixeira-Dias, *Molecular Physical Chemistry*,
DOI 10.1007/978-3-319-41093-7_2

2.1 Rate of a Chemical Reaction

Add bleach (typically a solution of sodium hypochlorite) to a food dye in aqueous solution. This simple experiment can illustrate the change of concentration of the dye as a function of time. Let us call the dye A and assume that its initial concentration is $[A]_0 = 5.0 \times 10^{-5}$ M. The continuous addition of bleach causes progressive discoloring. The spectral absorbance of such a diluted solution is proportional to the concentration of the absorbing species (Beer–Lambert law). This empirical, relationship makes it possible to determine the dye concentration from absorbance measurements taken at an absorption maximum wavelength (λ_{max}). Since the initial concentration of the dye is very low, the bleach exists in large excess as compared with the dye. Hence, the bleach concentration does not appreciably change during dye discoloring and the chemical reaction can be represented in the following way

$$A \rightarrow B \tag{2.1}$$

where B represents the colorless form of the dye (the bleach is omitted because its concentration is approximately constant during the experiment). The following table gives the concentration of A as a function of time at 1-min intervals.

t/min	0	1	2	3	4	5	6	7	8	9	10
$[A]/10^{-5}$ M	5.00	3.57	2.49	1.82	1.26	0.89	0.61	0.45	0.30	0.23	0.16

The graph of [A] as a function of t is called a **kinetic reaction profile** for the chemical reaction (2.1) and is presented in Fig. 2.1 (see *Mathematica* code **M1**). Once the fitting curve is known (an exponential, as we will soon find out), it is possible to obtain the derivative at each instant in the considered time range. Since A is the reactant, [A] is a decreasing function of time, its first derivative is negative, and the **reaction rate** is given by $-d[A]/dt$. The reaction rate defined using B is given by $d[B]/dt$, that is, $d[B]/dt = -d[A]/dt$ (**E1**).

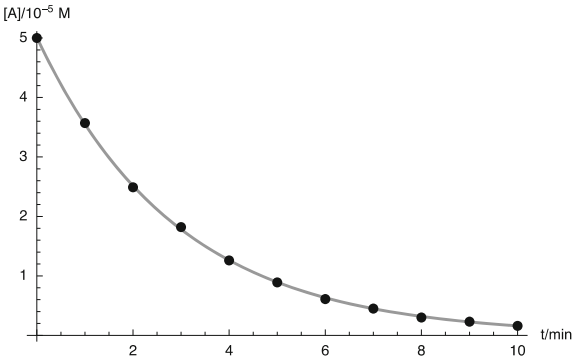


Fig. 2.1 Typical kinetic reaction profile for discoloration of a dye. Figure obtained with *Mathematica*

Considering now the chemical reaction

$$\mathrm{A} + \mathrm{B} \rightarrow 2\mathrm{C} \tag{2.2}$$

[for example, $H_2(g) + I_2(g) \rightarrow 2HI(g)$], we can write

$$\mathrm{d}n_\mathrm{C} = 2(-\mathrm{d}n_\mathrm{A}) = 2(-\mathrm{d}n_\mathrm{B}) \tag{2.3}$$

where $\mathrm{d}n_\mathrm{A}$, $\mathrm{d}n_\mathrm{B}$, and $\mathrm{d}n_\mathrm{C}$ represent changes in the molar quantities of A, B, and C. Both $\mathrm{d}n_\mathrm{A}$ and $\mathrm{d}n_\mathrm{B}$ are negative, because A and B are reactants, $\mathrm{d}n_\mathrm{C}$ is positive (C is the single product) and is twice $-\mathrm{d}n_\mathrm{A}$ or $-\mathrm{d}n_\mathrm{B}$ [see (2.2)]. From (2.3), we can write

$$\frac{\mathrm{d}n_\mathrm{A}}{-1} = \frac{\mathrm{d}n_\mathrm{B}}{-1} = \frac{\mathrm{d}n_\mathrm{C}}{2} \tag{2.4}$$

where the denominators are the stoichiometric numbers (=stoichiometric coefficients, where the stoichiometric coefficients of reactants are multiplied by −1). For the generic chemical reaction

$$a\mathrm{A} + b\mathrm{B} \rightarrow c\mathrm{C} + d\mathrm{D} \tag{2.5}$$

we can write

$$\frac{\mathrm{d}n_\mathrm{A}}{-a} = \frac{\mathrm{d}n_\mathrm{B}}{-b} = \frac{\mathrm{d}n_\mathrm{C}}{c} = \frac{\mathrm{d}n_\mathrm{D}}{d} \equiv \mathrm{d}\xi \tag{2.6}$$

where ξ is called the **extent** of the chemical reaction that specifies the variable composition along the course of the chemical reaction (ξ is a state variable). If the chemical reaction (2.5) is carried out at constant volume, then

$$\frac{1}{-a}\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} = \frac{1}{-b}\frac{\mathrm{d}[\mathrm{B}]}{\mathrm{d}t} = \frac{1}{c}\frac{\mathrm{d}[\mathrm{C}]}{\mathrm{d}t} = \frac{1}{d}\frac{\mathrm{d}[\mathrm{D}]}{\mathrm{d}t} = \frac{1}{V}\frac{\mathrm{d}\xi}{\mathrm{d}t} \tag{2.7}$$

where the square brackets stand for molarity. Each member of (2.7) defines the rate of the chemical reaction that can be obtained at any specified instant during the course of the chemical reaction (**E2**). Note that (2.7) assumes a time-independent stoichiometry, that is, the reaction stoichiometry is assumed to be valid at every instant during the reaction. This assumption is not always true. If the chemical reaction has intermediate species with lifetimes that enable their detection, then the stoichiometry is not time-independent, and the rate of the reaction cannot be defined in a unique way. In these cases, we are really talking about a sequence of chemical reactions, and ξ is a state variable for *each* chemical reaction in the sequence.

2.2 Experimental Rate Equation

Experimental studies show that the rates of many chemical reactions depend on the concentrations of the reactants according to the following equation:

$$v = k[\mathrm{A}]^x[\mathrm{B}]^y \tag{2.8}$$

where A and B are reactants, the exponents x and y are called **partial orders** of A and B, and k is called the **rate constant**. The **overall order** of the reaction is equal to $x + y$. Since k, x, y, and the concentrations in the second member of (2.8) are experimentally determined quantities, (2.8) is called the **experimental rate equation**. Partial orders x and y usually take the values 1, 2, or 0 (when the rate equation does not depend on the concentration of a particular reactant, its partial order with respect to that reactant is zero). Since the reaction rate depends on temperature, the experimental rate equation is *always* determined at a specified temperature.

Most chemical reactions have more than one reactant. When the partial order of one particular reactant is to be experimentally determined, conditions should be set to guarantee that other reactants do not interfere, and their concentrations will not vary significantly. For instance, considering a chemical reaction with reactants A and B, the experimental determination of the partial order of reactant A requires that [B] be kept approximately constant. Under these conditions, (2.8) takes the form

$$v = k'[\mathrm{A}]^x \tag{2.9}$$

where $k' = k[\mathrm{B}]^y \approx$ constant, because [B] does not significantly change in the course of the reaction. Equation (2.9) is called the **pseudo rate equation**, and k' represents the **pseudo rate constant**. The partial order x acts like the overall pseudo-order for the chemical reaction. The isolation and initial rate methods allow separation of the concentration variables of a rate equation in order to determine the partial order of a particular reactant.

In the experimental determination of the partial order of A by the **isolation method**, the concentration of B should greatly exceed the initial concentration of A. If A and B are in the stoichiometric proportion 1:1 and the initial concentrations of A and B are in the ratio 1:100 (for example, $[\mathrm{A}]_0 = 0.100$ M and $[\mathrm{B}]_0 = 10.0$ M), then when 99 % of A has reacted, the change in the concentration of B is only 1 %.

The **initial rate method** takes the rate equation at the initial instant ($t = 0$),

$$v_0 = k[\mathrm{A}]_0^x[\mathrm{B}]_0^y \tag{2.10}$$

The determination of the partial order of A requires kinetic experiments with different initial concentrations of A for the same initial concentration of B. For chemical reactions with products that decompose or interfere during the course of the reaction, the initial rate method is the sole method available for the kinetic study of such chemical reactions.

2.2.1 First-Order Reactions

The rate equation for a **first-order chemical reaction** with a single reactant (or a pseudo first-order chemical reaction) is given by

$$-\mathrm{d}[\mathrm{A}]/\mathrm{d}t = k[\mathrm{A}] \tag{2.11}$$

Solving this differential equation consists in finding the function $f(t) = [\mathrm{A}]$ that satisfies (2.11) and the initial condition $f(0) = [\mathrm{A}]_0$, where $[\mathrm{A}]_0$ is the initial concentration of A. While it is easy to conclude that the function that satisfies (2.11) is an exponential [the derivative of $\exp(-kt)$ is $-k\exp(-kt)$], we solve (2.11) in a more general way that can be applied to second-order reactions. We begin by separating variables [A] and t in each member,

$$\mathrm{d}[\mathrm{A}]/[\mathrm{A}] = -k\mathrm{d}t \tag{2.12}$$

Now we obtain indefinite integrals of both members. The integration limits for the first member are $[\mathrm{A}]_0$ and [A], and for the second member, are the corresponding values of time, that is, 0 and t. We obtain

$$\ln\left([\mathrm{A}]/[\mathrm{A}]_0\right) = -kt \tag{2.13}$$

This equation is equivalent to

$$\ln[\mathrm{A}] = \ln[\mathrm{A}]_0 - kt \tag{2.14}$$

that is, the graph of ln[A] as a function of t gives a straight line whose intercept is $\ln[\mathrm{A}]_0$ and whose slope is equal to $-k$. If we now consider the time interval $t = t_{1/2}$ such that [A] is half its initial value, $[\mathrm{A}] = [\mathrm{A}]_0/2$, then substitution of these equalities in (2.13) leads to

$$t_{1/2} = \ln 2/k \tag{2.15}$$

where the time interval $t_{1/2}$ is called the **half-life**. Contrary to what the name might suggest, the half-life is not half of the time to reaction completion. Equality (2.15), valid for first-order chemical reactions, shows that the half-life does not depend on $[\mathrm{A}]_0$. This is an important result that applies *only* to first-order chemical reactions. Hence, when the half-life of a chemical reaction is independent of the initial concentration, we conclude that the chemical reaction has first-order kinetics (Fig. 2.2).

Equality (2.13) can be rewritten in the following equivalent way,

$$[\mathrm{A}] = [\mathrm{A}]_0 \mathrm{e}^{-kt} \tag{2.16}$$

which shows that [A] has **exponential decay** with time. Spontaneous decays of radioactive atomic nuclei are first-order processes. Adapting the above equalities to

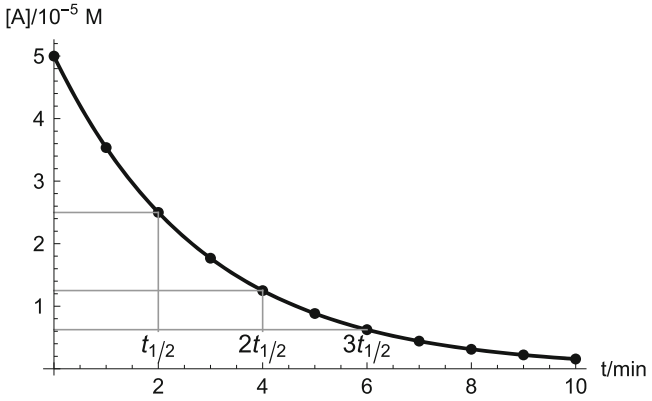


Fig. 2.2 The half-life of a first-order reaction does not depend on $[A]_0$. For the kinetic profile illustrated above, the half-life is approximately equal to 2 min. Figure obtained with *Mathematica*

radioactive decay simply requires substitution of [A] by the number of radioactive nuclei N and of $[A]_0$ by N_0. After these substitutions, Eq. (2.11) shows that the rate of radioactive decay ($-dN/dt$) is proportional to the number of nuclei that have not yet decayed (N). In radioactive studies, the half-life $t_{1/2}$ is preferred to the rate constant k, in contrast to what happens in chemical kinetics.

2.2.2 Second-Order Reactions

The rate equation for a second-order chemical reaction with a single reactant (or a pseudo second-order chemical reaction) is given by

$$-d[A]/dt = k[A]^2 \tag{2.17}$$

Separation of variables leads to

$$d[A]/[A]^2 = -k dt \tag{2.18}$$

After integration of both members, we obtain

$$\frac{1}{[A]} = \frac{1}{[A]_0} + kt \tag{2.19}$$

Hence, the graph of $1/[A]$ as a function of t gives a straight line whose intercept is $1/[A]_0$ and whose slope is equal to k. Substitution of [A] by $[A]_0/2$ and of t by $t_{1/2}$ in (2.19) gives the half-life for a second-order reaction:

$$t_{1/2} = \frac{1}{k[\mathrm{A}]_0} \tag{2.20}$$

Hence, the half-life of a second-order reaction is inversely proportional to the initial concentration.

2.2.3 Zeroth-Order Reactions

Zeroth order with respect to a particular reactant means that the chemical reaction rate does not depend on the concentration of that reactant. This can occur when the reactant is not involved in the slowest step of the reaction (the rate-limiting or **rate-determining step**), or when the reaction rate depends on the adsorption of a particular reactant by a fully covered catalyst surface. Since the surface is saturated, the reaction rate does not depend on the concentration of the adsorbed reactant.

The rate equation of a zeroth-order chemical reaction for a single reactant is given by

$$-\mathrm{d}[\mathrm{A}]/\mathrm{d}t = k[\mathrm{A}]^0 = k \tag{2.21}$$

After separation of variables and integration of both members, we obtain

$$[\mathrm{A}] = [\mathrm{A}]_0 - kt \tag{2.22}$$

Thus, the graph of [A] as a function of t gives a straight line whose intercept is $[\mathrm{A}]_0$ and whose slope is equal to $-k$. Substitution of [A] by $[\mathrm{A}]_0/2$ and of t by $t_{1/2}$ in (2.22) leads to

$$t_{1/2} = [\mathrm{A}]_0/(2k) \tag{2.23}$$

Hence, the half-life of a zeroth-order reaction is proportional to the initial concentration.

2.3 Effect of Temperature Change

The temperature-dependence of the rate of a chemical reaction lies essentially in the rate constant [see (2.8)], since the concentration factors expressed in molality are not affected by temperature (concentrations that involve a volume in their definition, like molarity, might be slightly affected by temperature variation). Based on observation of rate constant variations with temperature, Arrhenius (1859–1927; Nobel Prize in chemistry in 1903) proposed the $k(T)$ empirical, dependence given by

$$k(T) = A \exp\left(-\frac{E_a}{RT}\right) \tag{2.24}$$

where E_a is the **Arrhenius activation energy** and A is the **Arrhenius *A*-factor** (**E3**). The logarithm of (2.24) is

$$\ln k(T) = \ln A - \frac{E_a}{RT} \tag{2.25}$$

thus showing that $\ln k(T)$ depends linearly on $1/T$. The intercept of the resulting straight line is $\ln A$ and the slope is equal to $-E_a/R$. The derivative of (2.25) is given by

$$\frac{d\ln k(T)}{dT} = \frac{E_a}{RT^2} \tag{2.26}$$

This equality suggests a way for determining the activation energy from experimental data on $k(T)$, provided the chemical reaction rate constant follows an Arrhenius dependence.

Above the troposphere, the stratosphere ranges in altitude from about 11–50 km, with temperature varying approximately from −60 to −2 °C. The ozone layer is mainly found in the lower layer of the stratosphere, up to an altitude of approximately 30 km. As an example of the Arrhenius dependence, consider the second-order stratospheric chemical reaction

$$N + O_2 \rightarrow NO + O \tag{2.27}$$

for which $A = 1.5 \times 10^{-11}$ cm^3 molecule^{-1} s^{-1} and $E_a/R = 3600$ K (values taken from *Handbook of Chemistry and Physics*, 2011). The Arrhenius temperature-dependence of the rate constant of the above chemical reaction is illustrated in Fig. 2.3.

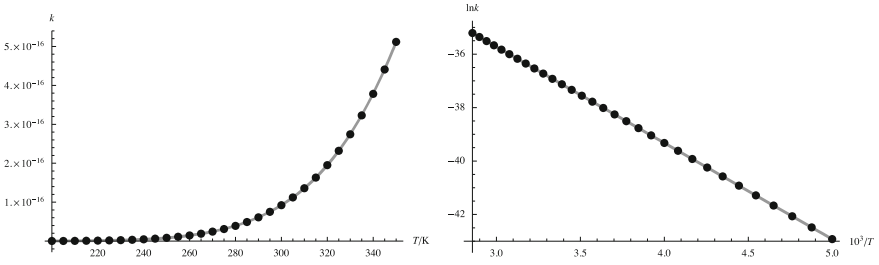


Fig. 2.3 Rate constant of the chemical reaction N + O_2 → NO + O as a function of temperature (*left*), and ln*k* as a function of $10^3/T$ (*right*), in the temperature range from 200 to 350 K. Note that the lowest value of $10^3/T$ in the horizontal axis of the plot at the right is far from the origin, which would correspond to an infinite temperature. Thus, the intersection of the straight line with the horizontal axis is far from the intercept. Graphs obtained with *Mathematica*

E_a/kJ mol^{-1}	exp[$-E_a/(RT)$]	
	300 K	600 K
10	2×10^{-2}	1×10^{-1}
50	2×10^{-9}	4×10^{-5}
100	4×10^{-18}	2×10^{-9}

The Arrhenius activation energy has a strong influence on the rate equation. To illustrate this point, the table shows approximate values of the Arrhenius exponential function for three hypothetical values of the activation energy (10, 50, and 100 kJ mol^{-1}), at two different temperatures (300 and 600 K). At 300 K, the Arrhenius exponential decreases by a factor of order 10^{-7} when the activation energy changes from 10 to 50 kJ mol^{-1}, and by a factor of order 10^{-9} when the activation energy changes from 50 to 100 kJ mol^{-1}. In turn, at 600 K, the same increments in the activation energy lead to exponential decreases by factors of order 10^{-4}. In conclusion, the decrease in the activation energy due to the eventual use of catalysts and the increase in temperature lead to appreciable increases in the chemical reaction rates that present the Arrhenius dependence. Note that the catalyst reduces the activation energy of the chemical reaction, thus leading to an increase in the reaction rate. The decrease in the activation energy is the result of a change of mechanism without altering the initial and final states of the reaction (the reactants and products of the overall chemical reaction).

In the graph of $\ln k$ as a function of $10^3/T$ [see (2.25)], the extrapolation needed to obtain the intercept may lead to a large uncertainty in the determination of the A-factor, since the data points are usually within a short range of $10^3/T$. For instance, the data points of Fig. 2.3 range from $10^3/T = 2.85$ K^{-1} ($T \approx 351$ K) to $10^3/T = 5.0$ K^{-1} ($T = 200$ K). At 1000 K, the value of $10^3/T$ is 1 K^{-1}, still far from the origin.

2.4 Elementary Reactions

Elementary reactions occur in a single step and so have time-independent stoichiometries and do not have reaction intermediates. An elementary reaction has a single potential energy maximum in the reaction path as a function of the reaction coordinate. In contrast, the existence of one reaction intermediate in a chemical reaction implies an energy minimum in the reaction path between reactants and products.

Elementary reactions can be classified according to their **molecularity**, which is the number of reactant molecules that take part in the reaction. Those reactions that involve one, two, and, less frequently, three reactant molecules are called **unimolecular**, **bimolecular**, and **trimolecular** reactions, respectively. There are few examples of chemical reactions thought to involve three reactant molecules, and no reactions are known to involve four reactant molecules, since it is highly

improbable that four molecules collide at the same instant with the energy and orientation required for a chemical reaction. Note that the term molecularity applies only to elementary reactions or to individual steps of complex reactions.

Consider a reaction of the following type:

$$Nu^- + RX \rightarrow RNu + X^- \tag{2.28}$$

where Nu^- stands for a nucleophile and X represents an electronegative atom or an electronegative group bonded to a tetrahedral carbon atom in the R radical. A chemical reaction of this type is called a **nucleophilic substitution reaction**. In fact, it is a substitution reaction (X is substituted by Nu) resulting from a nucleophilic attack of Nu^- on the carbon atom bonded to X:

$$Nu^- + C^{\delta+} - X^{\delta-} \tag{2.29}$$

If the nucleophilic substitution reaction occurs in a single step, the approach of Nu^- is *concerted* with the withdrawal of X^-. In this case, the transition state involves both reactant species Nu^- and RX, and the nucleophilic substitution reaction is bimolecular and named S_N2. The reaction is said to occur in a **bimolecular concerted step**, and its rate equation is given by

$$v = k[RX][Nu^-] \tag{2.30}$$

These considerations can be illustrated by the reaction of bromomethane with sodium hydroxide, in methanol, at 25 °C,

$$OH^- + CH_3Br \rightarrow CH_3OH + Br^- \tag{2.31}$$

This reaction has second-order kinetics and obeys the following experimental rate equation:

$$v = k[CH_3Br][OH]^- \tag{2.32}$$

This reaction mechanism has a single step with a single transition state, $[HO\ldots CH_3\ldots Br]^-$. The transition state involves both reactant species, and the nucleophilic substitution is bimolecular, that is, the mechanism is S_N2.

2.5 Complex Reactions

When the mechanism of a chemical reaction consists of more than one step, the reaction is said to be a **complex reaction**. Experimental evidence for the existence of one reaction intermediate leads to the conclusion that the reaction mechanism is formed by at least two steps. Sometimes, the lifetime of a reaction intermediate enables its isolation and characterization. However, reaction intermediates usually are reactive species with

very low concentrations, thus being difficult to detect. The physical methods generally used in the detection of short-lived reaction intermediates are spectroscopic methods, which involve an interaction with electromagnetic radiation.

The rate equation of an elementary reaction can be written once the reaction stoichiometry is known. Hence, when the rate equation does not reflect the reaction stoichiometry, we have a complex reaction. For example, the reaction between hypochlorite and iodide ions in aqueous solution,

$$ClO^- + I^- \rightarrow IO^- + Cl^- \tag{2.33}$$

has the following experimental rate equation:

$$v = k[ClO^-][I^-]/[OH^-] \tag{2.34}$$

The presence of the concentration of the hydroxide ion in the rate equation and its absence from the stoichiometric Eq. (2.33) leads to the conclusion that the chemical reaction (2.33) has more than one step in its mechanism; that is, it is a complex reaction.

The reaction of 2-bromo-2-methylpropane, $(CH_3)_3CBr$, with sodium hydroxide, in methanol at 25 °C,

$$OH^- + (CH_3)_3CBr \rightarrow (CH_3)_3COH + Br^- \tag{2.35}$$

has the following experimental first-order rate equation:

$$v = k[(CH_3)_3CBr] \tag{2.36}$$

This rate equation shows that the slowest reaction step involves a single reactant species, namely, $(CH_3)_3CBr$. Having the chemical reaction (2.35) in mind, we can infer from (2.36) that the first step of (2.35) consists in the dissociation of $(CH_3)_3CBr$ in the $(CH_3)_3C^+$ carbocation and the bromide ion,

$$(CH_3)_3CBr \rightleftarrows (CH_3)_3C^+ + Br^- \quad \text{slow step} \tag{2.37}$$

Considering the reactivity of the $(CH_3)_3C^+$ carbocation, it is likely that this chemical species recombines with the bromide ion to form 2-bromo-2-methylpropane, giving rise to an equilibrium. However, when the $(CH_3)_3C^+$ carbocation reacts with a hydroxide ion, an alcohol molecule is formed,

$$(CH_3)_3C^+ + OH^- \rightarrow (CH_3)_3COH \quad \text{fast step} \tag{2.38}$$

This is a fast step, because the hydroxide ion has a stronger nucleophilic character than the bromide ion. Addition of steps (2.37) and (2.38) brings back the original overall reaction (2.35). The transition state involves a single reactant molecule, $(CH_3)_3CBr$, and the overall reaction (2.35) represents a unimolecular nucleophilic substitution, thus being an example of an S_N1 mechanism.

The difference in the mechanisms of reactions (2.31) and (2.35) lies in the different stabilities of the involved carbocations, CH_3^+ and $(CH_3)_3C^+$. CH_3^+ is a stronger electrophile than $(CH_3)_3C^+$, because the same positive charge is distributed over a larger carbocation in $(CH_3)_3C^+$ as compared with CH_3^+. Hence, reaction (2.31) does not have a slow initial step for the dissociation of CH_3Br into CH_3^+ and Br^-, since CH_3^+ is quite a strong electrophile for that to occur.

2.6 Extremely Fast Reactions

Most reactions of ions in aqueous solution are extremely fast, reaching equilibrium in times of order 10^{-10} and 10^{-12} s. Such reactions cannot be studied using conventional methods that depend on the mixture of reactants, since the diffusion times (times for migration of reactant molecules until they collide with each other) are orders of magnitude greater than reaction times. Techniques can be used that apply an almost instantaneous perturbation, after which the system momentarily leaves the equilibrium state (concentrations of reactants and products are changed for a moment) and returns to the equilibrium state, since the duration of the perturbation is not significant when compared with the reaction half-life. The process of returning to the equilibrium state is called **relaxation**. The applied perturbation can be a shock wave, a pulse of electromagnetic radiation that produces a photochemical reversible reaction (flash photolysis), a sudden temperature jump (*T*-jump), or a sudden pressure jump associated with sound absorption in gaseous systems (Eigen 1954). The study of extremely fast reactions using brief energy pulses led to the award of the 1967 Nobel Prize in chemistry to Manfred Eigen, Ronald Norrish, and George Porter.

In order to proceed with the study of an extremely fast chemical reaction, it is necessary to know the rate equation for the reaction in the forward and reverse directions. Consider the reaction in aqueous solution

$$A + B \rightleftarrows C \tag{2.39}$$

where A and B are positive and negative ions and C is the chemical species resulting from the combination of those ions. Assume that the forward reaction is of first order in A and in B and that the reverse reaction is of first order in C, and represent by $k_{\rightarrow}$ and $k_{\leftarrow}$ the rate constants in the forward and reverse reactions. When a brief perturbation, for instance a temperature jump, is applied to the reaction at equilibrium, the A, B, and C concentrations will change with respect to their equilibrium values according to the following equalities, which preserve the reaction stoichiometry:

$$[A] = [A]_{eq} - \delta \quad [B] = [B]_{eq} - \delta \quad [C] = [C]_{eq} + \delta \tag{2.40}$$

where δ represents an infinitesimal change in concentration caused by the perturbation. These equalities can be written in the following equivalent way:

$$\Delta[\mathrm{A}] \equiv [\mathrm{A}] - [\mathrm{A}]_{\mathrm{eq}} = -\delta \quad \Delta[\mathrm{B}] \equiv [\mathrm{B}] - [\mathrm{B}]_{\mathrm{eq}} = -\delta \quad \Delta[\mathrm{C}] \equiv [\mathrm{C}] - [\mathrm{C}]_{\mathrm{eq}} = \delta \tag{2.41}$$

Therefore, the relaxation rate satisfies the following equalities:

$$-\frac{\mathrm{d}\Delta[\mathrm{A}]}{\mathrm{d}t} = -\frac{\mathrm{d}\Delta[\mathrm{B}]}{\mathrm{d}t} = \frac{\mathrm{d}\Delta[\mathrm{C}]}{\mathrm{d}t} = \frac{\mathrm{d}\delta}{\mathrm{d}t} \tag{2.42}$$

where, according to (2.39),

$$\frac{\mathrm{d}\delta}{\mathrm{d}t} = k_{\rightarrow}[\mathrm{A}][\mathrm{B}] - k_{\leftarrow}[\mathrm{C}] \tag{2.43}$$

Substitution of (2.40) in (2.43) leads to

$$\frac{\mathrm{d}\delta}{\mathrm{d}t} = \alpha - \beta\delta + \chi\delta^2 \tag{2.44}$$

where

$$\alpha = k_{\rightarrow}[\mathrm{A}]_{\mathrm{eq}}[\mathrm{B}]_{\mathrm{eq}} - k_{\leftarrow}[\mathrm{C}]_{\mathrm{eq}} \quad \beta = k_{\rightarrow}\left([\mathrm{A}]_{\mathrm{eq}} + [\mathrm{B}]_{\mathrm{eq}}\right) + k_{\leftarrow} \quad \chi = k_{\rightarrow} \tag{2.45}$$

Since the equilibrium constant of (2.39) is given by

$$K_{\mathrm{eq}} = \frac{k_{\rightarrow}}{k_{\leftarrow}} = \frac{[\mathrm{C}]_{\mathrm{eq}}}{[\mathrm{A}]_{\mathrm{eq}}[\mathrm{B}]_{\mathrm{eq}}} \tag{2.46}$$

it follows that $\alpha = 0$. In addition, in (2.44), the second-order term in δ is negligible when compared with the first-order term. Therefore,

$$\frac{\mathrm{d}\delta}{\mathrm{d}t} \approx -\beta\delta \tag{2.47}$$

This result shows that the perturbation δ follows first-order kinetics,

$$\delta = \delta_0 \mathrm{e}^{-\beta t} = \delta_0 \mathrm{e}^{-t/\tau} \tag{2.48}$$

where

$$\beta = \tau^{-1} \tag{2.49}$$

(the units of β are the inverse of time, s^{-1}) (**E4**). The rate constants $k_{\rightarrow}$ and $k_{\leftarrow}$ can be experimentally determined if two values of β, say β_1 and β_2, have been previously determined corresponding to two sets of $[A]_{eq}$ and $[B]_{eq}$ values (set_1 and set_2).

2.6.1 Neutralization Reaction in Water

Consider water and the neutralization equilibrium

$$H^+(aq) + OH^-(aq) \rightleftarrows H_2O(aq) \tag{2.50}$$

where $H^+(aq)$ stands for the symbolic representation of a protonated water molecule H_3O^+ or a protonated group of hydrogen-bonded water molecules as $H_5O_2^+$, $H_7O_3^+$, $H_9O_4^+$, …. The rate constant for the combination of H^+ and OH^- is equal to $k_{\rightarrow} = 1.4 \times 10^{11}$ $dm^3\ mol^{-1}\ s^{-1}$, that is, the neutralization is an extremely fast reaction, whereas the rate constant for the reverse reaction (ionization of H_2O) is equal to $k_{\leftarrow} = 2.5 \times 10^{-5}\ s^{-1}$ (Eigen 1967). At equilibrium, making use of (2.46), we can write

$$k_{\rightarrow}[H^+]_{eq}[OH^-]_{eq} - k_{\leftarrow}[H_2O] = 0 \tag{2.51}$$

where $[H_2O] \approx 55.6$ mol dm^{-3}. Hence, the constant for the ionic product of water at 25 °C is given by

$$[H^+]_{eq}[OH^-]_{eq} = \frac{k_{\leftarrow}}{k_{\rightarrow}}[H_2O] = 1.0 \times 10^{-14} \tag{2.52}$$

that is, $[H^+]_{eq} = [OH^-]_{eq} = 1.0 \times 10^{-7}$ mol dm^{-3}, at 25 °C.

2.7 Chemical Oscillations

Most chemical reactions function like one-way streets: the concentrations of reactants decrease, those of reaction intermediates increase at first and decrease later, and the concentrations of products increase. However, for a few reactions far from equilibrium, the concentrations of some chemical species oscillate, i.e., increase and decrease repeatedly. These reactions are called **chemical oscillators**. Chemical oscillations differ from pendulum oscillations: the film of a pendulum cannot be distinguished from the same film run backward (there is no arrow of time in the pendulum), whereas chemical oscillators are associated with entropy production due to irreversibleprocesses that occur in the reacting system.

The first reported observation of a periodic reaction in homogeneous solution is due to Bray in 1921 (Bray 1921). At that time, Lotka had reported the mathematical study of a system of differential equations describing the mechanism of a hypothetical periodic chemical reaction (Lotka 1920). The best-known oscillating chemical reaction results from experiments carried out by Belousov in 1958 and by Zhabotinsky in 1964, and is known as the **Belousov–Zhabotinsky experiment** (Winfree 1984). For years, the results of the Belousov–Zhabotinsky experiment were regarded with suspicion, since oscillations are incompatible with the existence of a Gibbs energy minimum at equilibrium. This apparent incompatibility was resolved when it was realized that chemical oscillations occur far from equilibrium.

2.7.1 Brusselator

In 1968, Prigogine and Lefever developed a model that shows how a chemical reaction, far from equilibrium, can pass from a stationary point to an oscillatory state (Prigogine 1968). This chemical oscillator, often called the **Brusselator** as a reminder of the Brussels thermodynamic school founded by Prigogine, consists of the following mechanism:

$$\begin{aligned} &\mathrm{A} \xrightarrow{k_1} \mathrm{X} \\ &\mathrm{B} + \mathrm{X} \xrightarrow{k_2} \mathrm{Y} + \mathrm{D} \\ &2\mathrm{X} + \mathrm{Y} \xrightarrow{k_3} 3\mathrm{X} \\ &\mathrm{X} \xrightarrow{k_4} \mathrm{E} \end{aligned} \tag{2.53}$$

where the inverse reactions are assumed to have negligible rate constants, and the overall chemical reaction is

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{D} + \mathrm{E} \tag{2.54}$$

In order to simplify the notation, *A, B, X,* and *Y* stand for [A], [B], [X], and [Y], respectively. Note that an increase of *X*, in elementary reactions 1 and 3, is followed by its decrease, in chemical reactions 2 and 4. A similar oscillatory behavior can be assigned to *Y* that decreases in chemical reaction 3, where Y is a reactant, and increases in reaction 2, where Y is a product.

The rate equations for the reaction intermediates X and Y in (2.53) are given by the following set of nonlinear coupled differential equations:

$$\frac{dX}{\mathrm{d}t} = k_1 A - k_2 BX + k_3 X^2 Y - k_4 X \quad \frac{dY}{\mathrm{d}t} = k_2 BX - k_3 X^2 Y \tag{2.55}$$

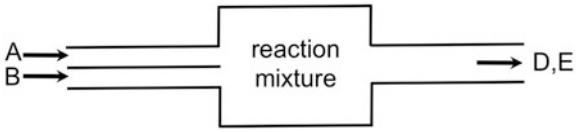


Fig. 2.4 In a flow reactor, the constant supply of A and B keeps the concentrations of these reactants in the reaction mixture fixed, and the constant removal of D and E maintains the system far from these reactants equilibrium

where the effects of diffusion of X and Y have been ignored, since we assume that chemical reactions (2.53) occur in a homogeneous medium, and so there is spatial uniformity of concentrations of all involved chemical species. In solving this set of nonlinear differential equations, the initial conditions are such that the corresponding physical system (the reaction mixture) is kept far from equilibrium: A and B have fixed values, while D and E are constantly removed from the reaction mixture to prevent reverse chemical reactions from occurring (Fig 2.4). Solving (2.55) for its stationary-point solutions ($\mathrm{d}X/\mathrm{d}t = 0$ and $\mathrm{d}Y/\mathrm{d}t = 0$) leads to

$$X_s = \frac{k_1}{k_4}A \quad Y_s = \frac{k_2 k_4}{k_1 k_3}\frac{B}{A} \tag{2.56}$$

where the subscript s stands for **stationary point**.

In order to assess the linear stability of the stationary point, we consider a time-dependent perturbation added to the stationary point solutions (2.56) and find out whether the perturbation increases or decreases in time (see *Kondepudi and Prigogine*, Further Reading).

The rate Eq. (2.55) can be written in the following more general way:

$$\frac{dX}{\mathrm{d}t} = Z_1(X, Y) \quad \frac{dY}{\mathrm{d}t} = Z_2(X, Y) \tag{2.57}$$

The time-dependent perturbations $x(t)$ and $y(t)$ are added to the stationary-point coordinates X_s and Y_s,

$$X = X_s + x(t) \quad Y = Y_s + y(t) \tag{2.58}$$

Therefore,

$$\frac{dX}{\mathrm{d}t} = \frac{dx}{\mathrm{d}t} \quad \frac{dY}{\mathrm{d}t} = \frac{dy}{\mathrm{d}t} \tag{2.59}$$

since X_s and Y_s are constants of the experiment. The Taylor expansion of Z_1 and Z_2 in (2.57) about the stationary point leads to

$$\begin{aligned} Z_1(X_s + x, Y_s + y) &= Z_1(X_s, Y_s) + \left(\frac{\partial Z_1}{\partial X}\right)_s x + \left(\frac{\partial Z_1}{\partial Y}\right)_s y + \cdots \\ Z_2(X_s + x, Y_s + y) &= Z_2(X_s, Y_s) + \left(\frac{\partial Z_2}{\partial X}\right)_s x + \left(\frac{\partial Z_2}{\partial Y}\right)_s y + \cdots \end{aligned} \tag{2.60}$$

Since $x(t)$ and $y(t)$ represent small perturbations, only the linear terms in (2.60) are retained (quadratic and higher-order terms are ignored). Hence, making use of (2.57) and (2.59), and noting that $Z_1(X_s, Y_s)$ and $Z_2(X_s, Y_s)$ are zero by definition of the stationary point, we conclude that

$$\frac{dx}{dt} = \left(\frac{\partial Z_1}{\partial X}\right)_s x + \left(\frac{\partial Z_1}{\partial Y}\right)_s y \quad \frac{dy}{dt} = \left(\frac{\partial Z_2}{\partial X}\right)_s x + \left(\frac{\partial Z_2}{\partial Y}\right)_s y \tag{2.61}$$

or in vector notation,

$$\frac{d\boldsymbol{x}}{dt} = \Lambda \boldsymbol{x} \tag{2.62}$$

where

$$\Lambda = \begin{bmatrix} \left(\frac{\partial Z_1}{\partial X}\right)_s & \left(\frac{\partial Z_1}{\partial Y}\right)_s \\ \left(\frac{\partial Z_2}{\partial X}\right)_s & \left(\frac{\partial Z_2}{\partial Y}\right)_s \end{bmatrix} \quad \boldsymbol{x} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \tag{2.63}$$

with Λ being the **Jacobian matrix**. The Jacobian matrix elements quantify the variation of the Z_1 and Z_2 rates with respect to changes in X and Y at the stationary point.

For the Brusselator, the Jacobian matrix is given by

$$\Lambda = \begin{pmatrix} k_2 B - k_4 & k_3 X_s^2 \\ -k_2 B & -k_3 X_s^2 \end{pmatrix} \tag{2.64}$$

where X_s is given by the first equality of (2.56). The **characteristic equation** (see **§1**) is

$$\det(\Lambda - \lambda \mathrm{I}) = 0 \tag{2.65}$$

that is,

$$\begin{vmatrix} k_2 B - k_4 - \lambda & k_3 X_s^2 \\ -k_2 B & -k_3 X_s^2 - \lambda \end{vmatrix} = 0 \quad \lambda^2 - (k_2 B - k_4 - k_3 X_s^2)\lambda + k_3 k_4 X_s^2 = 0 \tag{2.66}$$

The eigenvalues $\lambda_{\pm}$ are obtained by solving Eq. (2.66),

$$\lambda_{\pm} = \frac{p}{2} \pm \frac{1}{2}(p^2 - 4q)^{1/2} \tag{2.67}$$

where

$$p = k_2 B - k_4 - k_3 X_s^2 \quad q = k_3 k_4 X_s^2 \tag{2.68}$$

(**E5**).

Consider now the following set of parameters $k_1 = k_2 = k_3 = k_4 = 1.0$ and the initial condition $A = 1.0$, leading to $X_s = 1.0$, $Y_s = B$ [see (2.56)], $p = B - 2$, $q = 1$ [see (2.68)], and $p^2 - 4q = B^2 - 4B$. For $B = 2$, we have $p = 0$, and the eigenvalues of the Jacobian matrix are pure imaginary. For $B = 1.5$ or 2.5, we have $p^2 - 4q = (\pm 0.5)^2 - 4 < 0$, and the eigenvalues have a nonzero imaginary part. The value $B = 1.5$ leads to $p < 0$ and to a complex conjugate pair with a negative real part, whereas $B = 2.5$ gives $p > 0$ and a positive real part. We now return to the initial conditions $A = 1.0$ and $B = 2$. From (2.62) and (2.64), we can write

$$\begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \tag{2.69}$$

The eigenvalues of this Jacobian matrix are $\pm i$, and the corresponding eigenvectors are $(-1 - i, 2)^{\mathrm{T}}$ and $(-1 + i, 2)^{\mathrm{T}}$. Then the general solution to the above system of differential equations (see **§2**) is

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 \begin{pmatrix} -1 - i \\ 2 \end{pmatrix} \mathrm{e}^{it} + c_2 \begin{pmatrix} -1 + i \\ 2 \end{pmatrix} \mathrm{e}^{-it} \tag{2.70}$$

After applying $\mathrm{e}^{\pm it} = \cos t \pm i \sin t$, we obtain

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \cos t + \begin{pmatrix} \chi \\ \delta \end{pmatrix} \sin t \tag{2.71}$$

This equality shows that the $x(t)$ and $y(t)$ functions are linear combinations of cosine and sine functions. We now define U and V variables by $U \equiv X - X_s$ and $V \equiv Y - Y_s$, so that the stationary point is at $(U_s, V_s) = (0, 0)$ instead of at (X_s, Y_s). Figure 2.5 considers the same set of parameters ($k_1 = k_2 = k_3 = k_4 = 1.0$) and the initial condition $A = 1.0$, with B equal to 1.5, 2.0, and 2.5, and shows U and V as functions of time in the first plot and $V(t)$ as a function of $U(t)$ in the second plot (**phase trajectory**). For $B = 1.5$, the eigenvalues of the Jacobian matrix are a

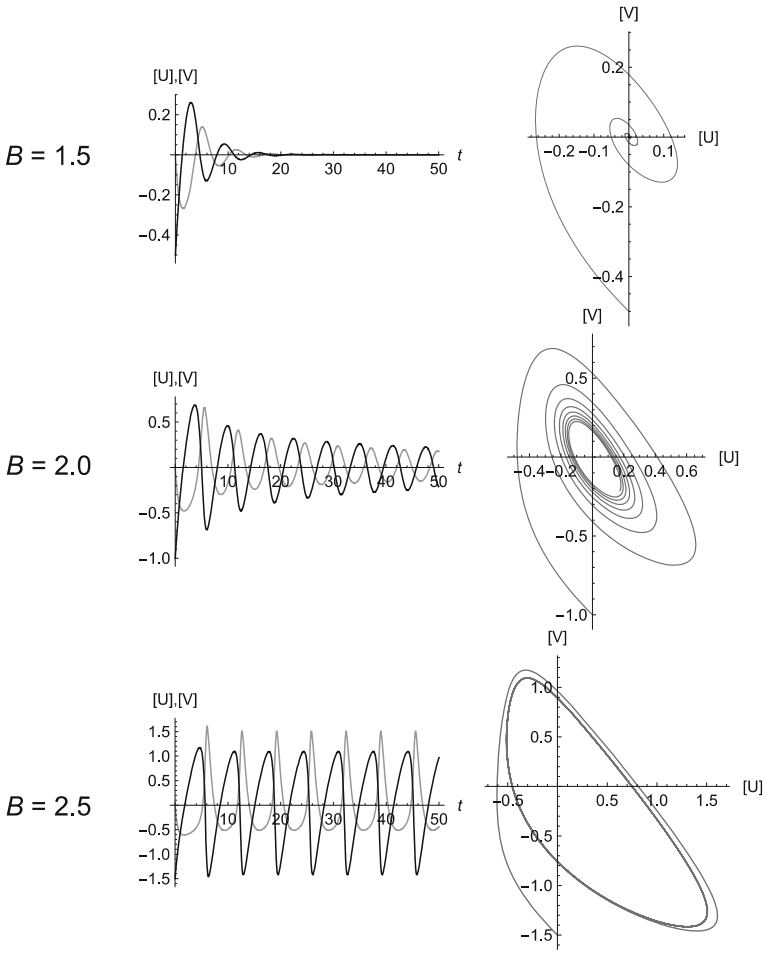


Fig. 2.5 Oscillations and phase trajectories for the Brusselator. Figures obtained with *Mathematica*

complex conjugate pair with a negative real part, and U and V quickly decay to zero as the phase trajectories rapidly spiral into the origin. For $B = 2.0$, the eigenvalues of Jacobian matrix are pure imaginary, and the trajectories spiral asymptotically to the origin as $t \to \infty$ (see *Mathematica* code **M2**). In turn, when $B = 2.5$, the eigenvalues of the Jacobian matrix are a complex conjugate pair with a positive real part, and the trajectories quickly approach a stable limit cycle as $t \to \infty$, with $U(t)$ and $V(t)$ indefinitely maintaining the same amplitudes and same shapes (the system is said to have reached permanent oscillations) (**E6**).

Notes

§1. Matrix Diagonalization

When a linear transformation represented by an $n \times n$ matrix A is applied to a vector $\boldsymbol{u}$ represented by an $n \times 1$ column vector, another $n \times 1$ vector $\boldsymbol{w}$ of the same vector space is obtained according to the following equation:

$$A\boldsymbol{u} = \boldsymbol{w} \tag{2.72}$$

This equality represents a linear transformation within the same vector space, which is said to preserve the vector length, since $\boldsymbol{u}$ and $\boldsymbol{w}$ have the same length. One especial and very important case of (2.72) quite frequent in quantum mechanics occurs when the vector $\boldsymbol{w}$ is equal to a scalar λ times the column vector $\boldsymbol{u}$, that is,

$$A\boldsymbol{u} = \lambda\boldsymbol{u} \tag{2.73}$$

where the scalar λ and the vector $\boldsymbol{u}$ are to be determined (they are not known beforehand). Certainly, the trivial solution $\boldsymbol{u} = \boldsymbol{0}$ is of no practical interest! Values of λ and nontrivial vectors $\boldsymbol{u}$ that satisfy (2.73) are called **eigenvalues** and **eigenvectors**, respectively. Equality (2.73) can be rewritten as

$$(A - \lambda I)\boldsymbol{u} = 0 \tag{2.74}$$

where the expression in parentheses is an $n \times n$ matrix. The equality (2.74) represents a homogeneous system of linear equations in n unknowns, and so the condition for a nontrivial solution of (2.74) is

$$\det(A - \lambda I) = |A - \lambda I| = 0 \tag{2.75}$$

Expansion of this determinant yields an nth-degree polynomial in λ. The polynomial Eq. (2.75) is called the **characteristic equation** or **secular equation**. The n roots of the characteristic equation are the n eigenvalues that satisfy (2.75). Substitution of each of these eigenvalues in (2.74) enables one to determine the corresponding eigenvector only to within a multiplicative constant, because if $\boldsymbol{u}$ is a solution of (2.74), so is any multiple of $\boldsymbol{u}$.

We can now illustrate the above considerations with the matrix

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix} \tag{2.76}$$

The characteristic equation is

$$\det(A-\lambda I)=\begin{vmatrix}1-\lambda & 0 & -1\\ 1 & 2-\lambda & 1\\ 2 & 2 & 3-\lambda\end{vmatrix}=-(\lambda-3)(\lambda-2)(\lambda-1)=0 \quad (2.77)$$

whose roots are $\lambda = 1, 2, 3$. We obtain the following eigenvectors:

$$\lambda=1\Rightarrow(A-I)\boldsymbol{u}_1=\begin{pmatrix}0 & 0 & -1\\ 1 & 1 & 1\\ 2 & 2 & 2\end{pmatrix}\boldsymbol{u}_1=0\Rightarrow\boldsymbol{u}_1=a\begin{pmatrix}-1\\ 1\\ 0\end{pmatrix} \quad (2.78)$$

$$\lambda=2\Rightarrow(A-2I)\boldsymbol{u}_2=\begin{pmatrix}-1 & 0 & -1\\ 1 & 0 & 1\\ 2 & 2 & 1\end{pmatrix}\boldsymbol{u}_2=0\Rightarrow\boldsymbol{u}_2=b\begin{pmatrix}-2\\ 1\\ 2\end{pmatrix} \quad (2.79)$$

$$\lambda=3\Rightarrow(A-3I)\boldsymbol{u}_3=\begin{pmatrix}-2 & 0 & -1\\ 1 & -1 & 1\\ 2 & 2 & 0\end{pmatrix}\boldsymbol{u}_3=0\Rightarrow\boldsymbol{u}_3=c\begin{pmatrix}-1\\ 1\\ 2\end{pmatrix} \quad (2.80)$$

The multiplicative constants a, b, and c can be determined once the lengths of $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_3$ are known. The eigenvectors $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_3$ can be grouped as columns of a 3×3 matrix $\boldsymbol{U}$, and the eigenvalues can be arranged as diagonal entries of the diagonal matrix D,

$$\boldsymbol{U}=\begin{pmatrix}| & | & |\\ \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3\\ | & | & |\end{pmatrix}\quad D=\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{pmatrix} \quad (2.81)$$

Using (2.74) for each of the eigenvalues λ_1, λ_2, and λ_3, we can write

$$\begin{aligned}A\boldsymbol{U}=&(A\boldsymbol{u}_1,A\boldsymbol{u}_2,A\boldsymbol{u}_3)=(\lambda_1\boldsymbol{u}_1,\lambda_2\boldsymbol{u}_2,\lambda_3\boldsymbol{u}_3)\begin{pmatrix}| & | & |\\ \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3\\ | & | & |\end{pmatrix}\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{pmatrix}\\ &=\boldsymbol{U}D\end{aligned} \quad (2.82)$$

Note that the product $\boldsymbol{U}D$ gives each eigenvalue corresponding to a column of the resulting matrix, as we want. On the other hand, the product $D\boldsymbol{U}$ would give each eigenvalue corresponding to a row of the product matrix. Considering (2.82), we can write

$$A\boldsymbol{U}=\boldsymbol{U}D \quad (2.83)$$

If the eigenvectors are linearly independent, then U is nonsingular, that is, its determinant is different from zero. Therefore, U^{-1} exists, and we can multiply (2.83) from the left by U^{-1} and obtain

$$\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U} = D \tag{2.84}$$

Two $n \times n$ matrices A and B are called **similar** if $B = X^{-1}AX$ for some nonsingular and therefore invertible matrix X. The transformation $A \rightarrow X^{-1}AX$ is a **similarity transformation**. Similar matrices represent the same transformation (operator) in two different bases of the same vector space related by the matrix X. Equality (2.84) shows a particular kind of a similarity transformation, whose result is a diagonal matrix. For that reason, (2.84) represents the **diagonalization** of the matrix A. The similarity transformation that diagonalizes the matrix A linearly combines the basis vectors, so that the resulting matrix becomes diagonal in the new basis.

We now illustrate the concepts of similarity transformation and similar matrices using a three-dimensional vector space with Cartesian coordinates and the basis vectors $(xyz)^{\mathrm{T}}$ and $(rst)^{\mathrm{T}}$. Consider the 3×3 matrix A that converts $(xyz)^{\mathrm{T}}$ into $(x'y'z')^{\mathrm{T}}$,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{2.85}$$

and the 3×3 matrix B that converts $(rst)^{\mathrm{T}}$ into $(r's't')^{\mathrm{T}}$,

$$\begin{pmatrix} r' \\ s' \\ t' \end{pmatrix} = B \begin{pmatrix} r \\ s \\ t \end{pmatrix} \tag{2.86}$$

Matrix X relates these bases: by applying the matrix X to $(rst)^{\mathrm{T}}$ and $(r's't')^{\mathrm{T}}$, we obtain $(xyz)^{\mathrm{T}}$ and $(x'y'z')^{\mathrm{T}}$, respectively,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = X \begin{pmatrix} r \\ s \\ t \end{pmatrix} \qquad \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = X \begin{pmatrix} r' \\ s' \\ t' \end{pmatrix} \tag{2.87}$$

The change of basis can be carried in the opposite direction, from $(xyz)^{\mathrm{T}}$ and $(x'y'z')^{\mathrm{T}}$ to $(rst)^{\mathrm{T}}$ and $(r's't')^{\mathrm{T}}$, meaning that X^{-1} exists, that is, X is nonsingular (its determinant is different from zero). If X is a symmetry operation, then as an element of an algebraic group, it has an inverse. Therefore, substitution of (2.87) in (2.85) gives

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Rightarrow X \begin{pmatrix} r' \\ s' \\ t' \end{pmatrix} = AX \begin{pmatrix} r \\ s \\ t \end{pmatrix} \Rightarrow \begin{pmatrix} r' \\ s' \\ t' \end{pmatrix} = X^{-1}AX \begin{pmatrix} r \\ s \\ t \end{pmatrix} \tag{2.88}$$

Comparison of this result with (2.86) leads to

$$B = X^{-1}AX \tag{2.89}$$

where $X^{-1}AX$ is called a **similarity transformation** and A and B are said to be **similar matrices**, that is, they represent the same linear transformation after a change of basis implemented by the matrix X. If both coordinate systems are orthogonal, then the inverse of matrix X is equal to its transpose, that is,

$$X^{-1} = X^{\mathrm{T}} \tag{2.90}$$

The matrix X is said to be **orthogonal**, and $X^{\mathrm{T}}AX$ is called an **orthogonal transformation**. An important conclusion about changing the vectorial basis within the same vector space is that a similarity transformation preserves the sum of the diagonal elements or trace of matrices A and B.

§2. Systems of First-Order Linear Differential Equations

Consider the system of two simultaneous first-order linear equations that corresponds to (2.62),

$$\dot{\boldsymbol{x}} = \Lambda \boldsymbol{x} \tag{2.91}$$

Since this system of differential equations has constant coefficients, it is reasonable to expect that a solution might be of the form

$$\boldsymbol{x}(t) = \boldsymbol{u}\mathrm{e}^{\lambda t} \tag{2.92}$$

where $\boldsymbol{u}$ and λ are to be determined. Substitution of (2.92) into (2.91) and cancellation of $\mathrm{e}^{\lambda t}$ from both sides gives the eigenvalue equation

$$\Lambda \boldsymbol{u} = \lambda \boldsymbol{u} \tag{2.93}$$

Since (2.91) is a system of *linear* differential equations, if $\boldsymbol{x}_1(t)$ and $\boldsymbol{x}_2(t)$ are solutions, then

$$\boldsymbol{x}(t) = c_1\boldsymbol{x}_1(t) + c_2\boldsymbol{x}_2(t) \tag{2.94}$$

is also a solution. Therefore, from (2.92) and (2.94) we conclude that the general solution for (2.91) is given by

$$\boldsymbol{x}(t) = c_1\boldsymbol{u}_1\mathrm{e}^{\lambda_1 t} + c_2\boldsymbol{u}_2\mathrm{e}^{\lambda_2 t} \tag{2.95}$$

where $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are eigenvectors and λ_1 and λ_2 are the corresponding eigenvalues (see *McQuarrie*, Sect. 11.6, p. 563).

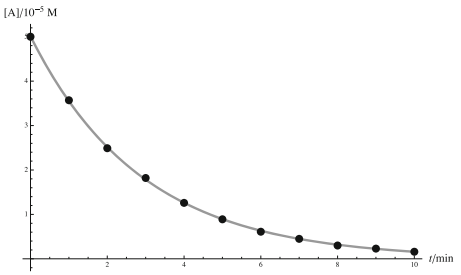
Mathematica Codes

M1. First-Order Chemical Reaction

```
data={{0.,5.00},{1.,3.57},{2.,2.49},{3.,1.82},{4.,1.26},{5.,0.89},{6.,0.61},{7.,0.45},
    {8.,0.30},{9.,0.23},{10.,0.16}};
model=a Exp[-k t];
fit=FindFit[data,model,{a,k},t]
f=Function[t,Evaluate[model/.fit]]
Plot[f[t],{t,0,10},PlotStyle→{Gray,Thickness[0.006]},
    AxesLabel→{Style["t/min",FontSize→14,Black],Style["[A]/10^-5 M",FontSize→14,Black]},
    Epilog→{PointSize[0.020],Black,Map[Point,data]},ImageSize→{550,350}]
```

```
{a → 5.01042, k → 0.344655}
```

```
Function[t, 5.01042 e^-0.344655 t]
```



This *Mathematica* code does a least squares fit of kinetic data to a model of exponential decay and plots the resulting exponential function and the kinetic data points. The first line of code presents a list of data points named `data`, and the second line defines the model of exponential decay

```
model=a Exp[-k t];
```

In turn, the third line of code,

```
fit=FindFit[data,model,{a,k},t]
```

does the fitting, where the command `FindFit` finds a least squares fit for the defined model function, and the second line of results presents the function

```
Function[t,5.01042 e^-0.344655t]
```

where the comma inside square brackets separates the time variable `t` from the body of the pure function

$$5.01042\,\mathrm{e}^{-0.344655\mathrm{t}}$$

The function `f[t]` is plotted in the last line of code, and the data points are rendered over the plotted curve by the *Mathematica* command `Epilog`, whose main action is

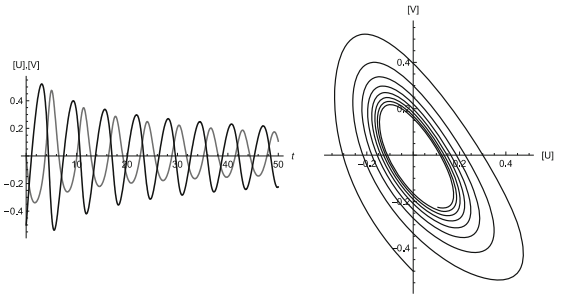
```
Map[Point,data]
```

This *Mathematica* function `Map` applies `Point` to each element of `data`.

Suggestion: Complete the code to evaluate the half-life to a precision of four decimal places.

M2. Brusselator

```
k1=1.0; k2=1.0; k3=1.0; k4=1.0;
A=1.0; B=2.0;
u0=0.0; v0=-0.5;
a=1.0; b=2.0;
mt=50; ms=10000;
sol=NDSolve[{U'[t]==k1 A-k2 B(U[t]+a)+k3 (V[t]+b)(U[t]+a)^2-k4 (U[t]+a),
    V'[t]==k2 B(U[t]+a)-k3 (V[t]+b)(U[t]+a)^2,U[0]==u0,V[0]==v0},
    {U,V},{t,0,mt},MaxSteps→ms];
Row[{Plot[Evaluate[{U[t],V[t]}/.sol],{t,0,mt},AxesLabel→{"t","[U],[V]"},AxesStyle→
    Directive[FontFamily→"Arial",FontSize→10],PlotStyle→{{Thickness[0.006],Darker[Gray]},
    {Thickness[0.006],Black}},PlotRange→All,ImageSize→{300,300}],
    ParametricPlot[Evaluate[{U[t],V[t]}/.sol],{t,0,mt},AxesLabel→{"[U]","[V]"},
    AxesStyle→Directive[FontFamily→"Arial",FontSize→10],
    PlotStyle→{Black,Thickness[0.006]},PlotRange→All,ImageSize→{300,300}]}]
```



This *Mathematica* code solves the equations of the Brusselator for $k_1 = k_2 = k_3 = k_4 = 1.0$, $A = 1.0$, and $B = 2.0$. Representing by $X(t)$ and $Y(t)$ the concentrations of the intermediates as functions of time and by X_s and Y_s the corresponding stationary values [for the above set of parameters, $X_s = 1.0$ and $Y_s = 2.0$, see (2.56)], the variables U and V are such that $X = U + X_s$ and $Y = V + Y_s$ and the rate equations are solved using the *Mathematica* command `NDSolve`, which finds a numerical solution for ordinary differential equations like the rate equations of the Brusselator [see (2.55)].

The last line of code plots in a row the $U(t)$ and $V(t)$ functions in the same graph and V as a function of U (a parametric plot) in a second graph. Removing from the last code line all the plot style options, we obtain

```
Row[{Plot[Evaluate[{U[t],V[t]}/.sol],{t,0,mt}],
ParametricPlot[Evaluate[{U[t],V[t]}/.sol],{t,0,mt}]}]
```

where the *Mathematica* function `Evaluate` causes `{U[t],V[t]}/.sol` to be evaluated, that is, `U[t]` and `V[t]` are replaced by the solutions of the rate equations `sol`. Note that the opposite of `Evaluate` is `Hold`, which maintains an expression to which it applies in an unevaluated form. `ParametricPlot` generates a plot of V as a function of U, since both of these variables are functions of time, an external variable, or parameter.

For the above-mentioned initial conditions ($k_1 = k_2 = k_3 = k_4 = 1.0$, $A = 1.0$, and $B = 2.0$), the eigenvalues of the Jacobian matrix are pure imaginary [$p = 0$ and

$q = 1$; see (2.67) and (2.68)], and the trajectories spiral asymptotically to the origin as $t \to \infty$, as can be concluded by inspection of the phase trajectory shown above.

Suggestion: Using the *Mathematica* function `Table`, write a *Mathematica* code for solving the equations of the Brusselator for $k_1 = k_2 = k_3 = k_4 = 1.0$, $A = 1.0$, and three values of B, namely, $B = 1.5$, 2.0, and 2.5.

Glossary

Arrhenius equation — The empirical exponential dependence of the rate constant as a function of temperature, proposed by Arrhenius (1859–1927; Nobel Prize in chemistry in 1903); see (2.24). Contains two empirical parameters, the Arrhenius A-factor and the Arrhenius activation energy E_a

Belousov–Zhabotinsky experiment — The best-known oscillating chemical reaction, resulting from experiments carried out by Belousov in 1958 and Zhabotinsky in 1964. For years, the results of this experiment were regarded with suspicion, since oscillations are incompatible with the existence of a Gibbs energy minimum at equilibrium. This apparent incompatibility was solved when it was realized that chemical oscillations occur far from equilibrium.

Brusselator — A model chemical oscillator developed in 1968 by Prigogine and Lefever, in the Brussels thermodynamic school founded by Prigogine, that shows how a chemical reaction, far from equilibrium, can pass from a stationary point to an oscillatory state; see (2.53).

Chemical oscillator — A complex reaction in which, far from equilibrium, the concentrations of some chemical species oscillate, i.e., increase and decrease repeatedly.

Complex reaction — A chemical reaction whose mechanism consists of more than one step, with the slowest step being the rate-determining step. When existing experimental evidence points to the occurrence of one reaction intermediate, then one can conclude that the reaction mechanism is formed by at least two steps.

Elementary reaction A chemical reaction that occurs in a single step, has a time-independent stoichiometry, and does not have any reaction intermediate. An elementary reaction has a single potential energy maximum in the reaction path as a function of the reaction coordinate.

Extremely fast reaction A chemical reaction that reaches equilibrium in times of order 10^{-10} and 10^{-12} s and so cannot be studied using conventional methods that depend on the mixture of reactants, since the diffusion times (times for migration of reactant molecules until they collide with each other) are orders of magnitude greater than the above reaction times.

Half-life The time interval required for the concentration of a reactant or the number of radioactive atoms to decrease to half its initial value.

Initial rate method Involves measuring the reaction rate of a chemical reaction at very short times before any significant changes in the concentrations of reactants occur; see (2.10).

Isolation method Involves measuring the reaction rate of a chemical reaction when the concentration of one reactant is greatly exceeded by the concentrations of all other reactants so that these do not significantly change during the reaction. Under these conditions, the rate equation takes a much simpler form with a pseudo rate constant; see (2.9) and compare with (2.8).

Kinetic reaction profile The graphical representation of the concentration of a reactant or product of a chemical reaction as a function of time.

Overall reaction order The sum of all partial orders in the experimental rate equation.

Partial order The exponent to which the concentration of a reactant is raised in the experimental rate equation. For many chemical reactions, the partial orders are not equal to the reaction stoichiometric coefficients, whereas for elementary chemical reactions, partial orders coincide with the stoichiometric coefficients.

Rate of chemical reaction The time derivative of the concentration of a reactant or product of a chemical reaction divided by the corresponding stoichiometric number; see (2.7).

Exercises

E1. Considering the experimental data that led to Fig. 2.1, use *Mathematica* to calculate the rate constant for the first-order chemical reaction and the time elapsed until the concentration of A is reduced to 5 % of its initial value.
E2. Derive (2.7).
E3. For a second-order chemical reaction and the following rate constant values, use *Mathematica* for determining the Arrhenius activation energy.

T/K	900	950	1000	1050	1100	1150
$k/M^{-1}\ s^{-1}$	0.01305	0.07686	0.37907	1.60593	5.96669	19.7777

E4. Derive (2.44).
E5. Consider the Brusselator.

(a) Write the set of differential equations for the concentrations of X and Y as functions of time and determine the stationary point.
(b) Obtain the Jacobian matrix at the stationary point. Assume $k_1 = k_2 = k_3 = k_4 = 1.0$ and $A = 1.0$, and determine the Jacobian matrix in terms of B.
(c) Use *Mathematica* for determining the eigenvalues of the Jacobian matrix for $B = 2.0$.

E6. Consider the Lotka–Volterra mechanism

$$\mathrm{A} + \mathrm{X} \xrightarrow{k_1} 2\mathrm{X}$$
$$\mathrm{X} + \mathrm{Y} \xrightarrow{k_2} 2\mathrm{Y}$$
$$\mathrm{Y} \xrightarrow{k_3} \mathrm{B}$$

(a) Write the set of differential equations for the concentrations of X and Y as functions of time and determine the stationary point.
(b) Determine the Jacobian matrix at the stationary point. Assume $k_1 = k_2 = k_3 = 1$ and $A = 1.0$ and determine the Jacobian matrix for these values.
(c) Use *Mathematica* for determining the eigenvalues of the Jacobian matrix.
(d) Use *Mathematica* for solving the kinetic equations, plot $U(t)$, $V(t)$, and the phase trajectory.

References

Bray WC (1921) A periodic reaction in homogeneous solution and its relation to catalysis. J Am Chem Soc 43:1262–1267

Eigen M (1954) Methods for investigation of ionic reactions in aqueous solutions with half-times as short as 10^{-9} s. Application to neutralization and hydrolysis reactions Discuss Faraday Soc 17:194–205

Eigen M (1967) Immeasurably fast reactions, Nobel Lecture, December 11

Lotka AJ (1920) Undamped oscillations derived from the law of mass action. J Am Chem Soc 42:1595–1599

Prigogine I, Lefever R (1968) Symmetry breaking instabilities in dissipative systems II. J Chem Phys 48:1695–1700

Winfree AT (1984) The prehistory of the Belousov-Zhabotinsky oscillator. J Chem Educ 61:661–663

Further Reading

Kondepudi D, Prigogine I (1998) Modern thermodynamics: from heat engines to dissipative structures. Wiley

McQuarrie D (2003) Mathematical methods for scientists and engineers. University Science Books

Mortimer M, Taylor P (eds) (2002) Chemical kinetics and mechanism. Series The molecular world. The Open University

The Schrödinger Equation

3

Abstract

The equation of motion of macroscopic particles was proposed by Isaac Newton toward the end of the seventeenth century. At the beginning of the twentieth century, several scientific advances led to the conclusion that classical mechanics was unable to correctly describe the mechanical behavior of subatomic particles. In particular, the electron, with its small mass of about 1 over 2×10^3 the proton mass, exhibits both particlelike and wavelike properties. This chapter aims at providing the student with the important background theory on quantum chemistry. To this end, it contains sections on operators, the harmonic oscillator, spherical coordinates, angular momentum, the hydrogen atom, the antisymmetry principle, the variational method, Born–Oppenheimer approximation, the Hartree–Fock method, and density functional theory, followed by sections on perturbation theory, time-dependent perturbation theory, absorption and emission of radiation and Raman scattering, and a final section on molecular calculations that comprises computational methods and standard basis sets. At the end of this chapter, the student finds two notes, one on the particle in a one-dimensional box and the other on the two-particle rigid rotor, that complement the main text without diverting the student's attention from the sequence of the chapter sections. Next, several *Mathematica* codes (Wave Equation, Helmholtz Equation, Harmonic Oscillator, Spherical Harmonics, Determinants, Systems of Homogeneous Linear Equations, Normalization Constants for Slater-Type Orbitals, Functional Derivative, STO versus a Gaussian Function at the Origin, Fitting Gaussian Functions to a 1s Hydrogen Orbital, Product of Gaussian Functions) are interconnected with the main text, along with detailed explanations of new *Mathematica* commands. Some of these have suggestions for the student to follow. Finally, the student will find a glossary of important scientific and technical terms and a list of exercises, whose complete answers are presented in a section at the end of the book.

J.J.C. Teixeira-Dias, *Molecular Physical Chemistry*,
DOI 10.1007/978-3-319-41093-7_3

3.1 Operators

An **operator** is a rule that transforms a given function into another function. For example, the derivative of $f(x)$ results from applying the operator $\mathrm{d}/\mathrm{d}x$ to $f(x)$,

$$f'(x) = \left(\frac{\mathrm{d}}{\mathrm{d}x}\right) f(x) \tag{3.1}$$

Two operators $\hat{A}$ and $\hat{B}$ are said to be equal if

$$\hat{A}f = \hat{B}f \tag{3.2}$$

for all functions f. The **composition** of two operators is defined by

$$\hat{A}\hat{B}f \equiv \hat{A}\left(\hat{B}f\right) \tag{3.3}$$

where the operator on the right is first applied to f, and the operator on the left is applied to the function in parentheses. The composition of operators is associative:

$$\hat{A}\left(\hat{B}\hat{C}\right) = \left(\hat{A}\hat{B}\right)\hat{C} \tag{3.4}$$

However, the composition of operators is not necessarily commutative, unlike constants and functions, whose product is always commutative. The **commutator** of two operators is defined by

$$\left[\hat{A}, \hat{B}\right] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \tag{3.5}$$

For two **commuting operators**, their commutator is equal to the null operator, and the order of application of both operators is irrelevant. The operator $\mathrm{d}/\mathrm{d}x$ and a constant c commute,

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) cf(x) = c\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) f(x) \Rightarrow \left[\frac{\mathrm{d}}{\mathrm{d}x}, c\right] = \left(\frac{\mathrm{d}}{\mathrm{d}x}\right) c - c\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) = 0 \tag{3.6}$$

but the operators $\mathrm{d}/\mathrm{d}x$ and x do not commute,

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)[xf(x)] = f(x) + x\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) f(x) \Rightarrow \left[\frac{\mathrm{d}}{\mathrm{d}x}, x\right] = \left(\frac{\mathrm{d}}{\mathrm{d}x}\right) x - x\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) = 1 \tag{3.7}$$

A **linear operator** satisfies the following equalities:

$$\hat{A}[f(x) + g(x)] = \hat{A}f(x) + \hat{A}g(x) \quad \hat{A}[cf(x)] = c\hat{A}f(x) \tag{3.8}$$

where f and g are arbitrary functions and c is an arbitrary constant. Examples of linear operators are x^2, $\mathrm{d}/\mathrm{d}x$, $\mathrm{d}^2/\mathrm{d}x^2$. Operators like cos and log are not linear.

3.1.1 Eigenvalues and Eigenfunctions

When the result of applying a linear operator to a function f is a constant times the same function,

$$\hat{A}f = af \tag{3.9}$$

we say that f is an **eigenfunction**, a is an **eigenvalue**, and (3.9) is the **eigenvalue equation** (*eigen*, from German adjective meaning *characteristic*). For example, e^{3x} is an eigenfunction of the operator $\mathrm{d}/\mathrm{d}x$ with eigenvalue equal to 3,

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)\mathrm{e}^{3x} = 3\mathrm{e}^{3x} \tag{3.10}$$

(**E1**, **E2**).

3.1.2 One-Dimensional Schrödinger Equation

The classical form of the kinetic energy in one dimension x is

$$\frac{p_x^2}{2m} \tag{3.11}$$

where p_x is the **linear momentum** along the x-axis, whose corresponding quantum mechanical operator is given by

$$\hat{p}_x = \frac{\hbar}{i}\frac{\partial}{\partial x} \tag{3.12}$$

(**E3**, **E4**).

The time-independent Schrödinger equation is given by the eigenvalue equation

$$\hat{H}\psi_k = E_k\psi_k \tag{3.13}$$

where the operator is called the **Hamiltonian** (the energy operator), ψ_k is the eigenfunction, and E_k is the corresponding energy eigenvalue. For a system consisting of one particle in one dimension, the Hamiltonian is given by

$$\hat{H} = \frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x) \tag{3.14}$$

where the first term is the kinetic energy operator (**E5**), the second term is the potential energy operator, and

$$\hbar = \frac{h}{2\pi} \tag{3.15}$$

Multiplication of (3.13) by the complex conjugate of ψ_k followed by integration over all coordinate space leads to

$$\int_{-\infty}^{+\infty} \psi_k^* \hat{H} \psi_k \mathrm{d}x = E_k \int_{-\infty}^{+\infty} \psi_k^* \psi_k \mathrm{d}x = E_k \tag{3.16}$$

where

$$\int_{-\infty}^{+\infty} \psi_k^* \psi_k \mathrm{d}x = \int_{-\infty}^{+\infty} |\psi_k|^2 \mathrm{d}x = 1 \tag{3.17}$$

The square of the absolute value of the wave function, $|\psi_k(x)|^2$, is the probability density for finding the particle at state ψ_k and point x, and the integral of $|\psi_k(x)|^2$ over the entire domain of x is 1, since ψ_k is normalized. The physical meaning of $|\psi_k|^2$ as a probability density implies that ψ_k is single-valued and continuous, has continuous first derivatives, and is quadratically integrable, that is, the integral of $|\psi_k|^2$ is finite and so can be normalized. A function that meets these requirements is called a **well-behaved function**.

The Schrödinger equation cannot be derived any more than the Newton equation of classical mechanics can be derived. In fact, these equations cannot be derived from first principles, since they express first principles for the mechanical behavior of macroscopic bodies and subatomic particles. In spite of that, we can imagine the line of thought followed by Schrödinger to arrive at his famous and less intuitive equation. To this end, the classical wave equation is combined with the De Broglie expression that connects linear momentum, a particle characteristic, with wavelength, a wave feature. We begin by writing the following **one-dimensional wave equation**,

$$\frac{\partial^2 u(x,t)}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 u(x,t)}{\partial t^2} \tag{3.18}$$

where u is any twice differentiable function, and v represents the speed of the propagating wave. The *Mathematica* code **M1** solves this differential equation, showing that its general solution is a linear combination of any twice differentiable functions of arguments $x - vt$ and $x + vt$. Functions with these arguments are called waves and propagate to the right or forward direction (argument $x - vt$) and to the left or backward direction (argument $x + vt$) (**E6**).

A propagating wave conveys two kinds of periodicity: *space* periodicity in the x direction with wavelength λ as its characteristic parameter, and *time* periodicity with

period τ, the inverse of frequency ν, as its characteristic parameter. For a forward wave, we can take $x/\lambda - t/\tau$, since both x/λ and t/τ are dimensionless quotients and the minus sign indicates a forward-propagating wave. A sine or cosine wave is called a harmonic wave. The units of the argument of a sine function are radians, so $x/\lambda - t/\tau$ should be multiplied by 2π, giving $2\pi\ (x/\lambda - t/\tau)$, where $2\pi\ /\lambda$ is the **angular wavenumber** k and $2\pi/\tau = 2\pi\nu$ is the **angular frequency** ω. Therefore, a sine forward wave is given by $A\sin(kx - \omega t)$, where A is the wave amplitude (**E7**).

Assuming separation of variables x and t and expressing the time-dependence of $u(x, t)$ by a cosine function, we can write

$$u(x,t) = \psi(x)\cos\omega t \tag{3.19}$$

Substitution of (3.19) in (3.18) gives

$$\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} + k^2\psi(x) = 0 \tag{3.20}$$

where $\nu = \nu\lambda$ and k is the angular wavenumber. The *Mathematica* code **M2** solves this differential equation, known as the **Helmholtz equation** (see also **§1**). We now consider the particle energy as a sum of the kinetic and potential energies,

$$E = \frac{p_x^2}{2m} + V(x) \tag{3.21}$$

and use the De Broglie equation

$$\lambda p_x = h \tag{3.22}$$

Substitution of (3.21) and (3.22) in (3.20) gives the **one-dimensional Schrödinger equation**

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} + V(x)\psi(x) = E\psi(x) \tag{3.23}$$

(**E8**, **E9**, **E10**).

3.1.3 Hermitian Operators

Consider now an arbitrary well-behaved function f and a set $g_1, g_2, \ldots, g_i, \ldots$ of well-behaved functions that obey the same boundary conditions of f and satisfy the equality

$$f = \sum_i c_i g_i \tag{3.24}$$

where the c_i are coefficients. The set of functions g_1, g_2, …, g_i, … is said to be a **complete set of functions** and can be used to expand any arbitrary function that obeys the same boundary conditions of the functions g_i.

Now consider a linear operator $\hat{A}$ that represents the physical quantity A. When the system is described by the state function ψ, the average value of $\hat{A}$ is given by

$$\langle A \rangle = \int \psi^* \hat{A} \psi \mathrm{d}\tau \tag{3.25}$$

where the integral is over the full range of coordinates of the wave function. The average value of a physical quantity is a real number, so

$$\langle A \rangle = \langle A \rangle^* \tag{3.26}$$

and

$$\int \psi^* \hat{A} \psi \mathrm{d}\tau = \int \psi \left(\hat{A}\psi\right)^* \mathrm{d}\tau \tag{3.27}$$

A linear operator that satisfies this equality for all well-behaved functions is called a **Hermitian operator**. Replacing ψ by one eigenfunction of the Hermitian operator, one easily concludes that the eigenvalues of a Hermitian operator are real numbers. The above definition of Hermitian operator can be shown to be equivalent to the following,

$$\int f^* \hat{A} g \mathrm{d}\tau = \int g \left(\hat{A} f\right)^* \mathrm{d}\tau \tag{3.28}$$

for arbitrary well-behaved functions f and g (**E11**).

3.1.4 Important Theorems

We are now going to prove a few important theorems in quantum mechanics.

Theorem 1 *The eigenvalues of a Hermitian operator are real numbers.*

Consider the Hermitian operator $\hat{A}$ and its general eigenvalue equation

$$\hat{A} g_k = a_k g_k \tag{3.29}$$

Using the definition of Hermitian operator, we write

$$\int g_k^* \hat{A} g_k \mathrm{d}\tau = \int g_k \left(\hat{A} g_k\right)^* \mathrm{d}\tau \tag{3.30}$$

Substitution of (3.29) in (3.30) leads to

$$\left(a_k - a_k^*\right) \int g_k^* g_k \mathrm{d}\tau = 0 \tag{3.31}$$

Since g_k is a quadratically integrable function, we have

$$a_k = a_k^* \tag{3.32}$$

that is, the eigenvalue a_k is real. Therefore, the eigenvalues of a Hermitian operator are real numbers.

Theorem 2 *Two linear operators with a common complete set of eigenfunctions commute.*

If two linear operators have a common complete set of eigenfunctions g_1, g_2, …, g_i, …,

$$\hat{A} g_i = a_i g_i \quad \hat{B} g_i = b_i g_i \tag{3.33}$$

where a_i and b_i are eigenvalues, then they commute, i.e., their commutator is the null operator, and consequently,

$$\left(\hat{A}\hat{B} - \hat{B}\hat{A}\right) f = 0 \tag{3.34}$$

for every arbitrary well-behaved function f. To arrive at (3.34), express f as a linear combination of the complete set of functions g_i [see (3.24)] and use these as eigenfunctions common to both operators [see (3.33)].

Theorem 3 *If two Hermitian operators commute, then there exists a common complete set of eigenfunctions for them.*

In order to demonstrate this theorem, we consider two Hermitian operators $\hat{A}$ and $\hat{B}$ and write the eigenvalue equation for $\hat{A}$,

$$\hat{A} g_i = a_i g_i \tag{3.35}$$

If we now apply $\hat{B}$ to both members, we obtain

$$\hat{B}\left(\hat{A} g_i\right) = \hat{B}(a_i g_i) = a_i \left(\hat{B} g_i\right) \tag{3.36}$$

Being Hermitian, $\hat{B}$ is linear, and this justifies the last equality. Since $\hat{A}$ and $\hat{B}$ commute, we can write

$$\hat{A}\left(\hat{B} g_i\right) = a_i \left(\hat{B} g_i\right) \tag{3.37}$$

that is, the function in parentheses is an eigenfunction of $\hat{A}$ with eigenvalue a_i. Assuming that the eigenvalues of $\hat{A}$ are nondegenerate (degenerate eigenvalues are equal), then each eigenvalue corresponds to one and only one linearly independent eigenfunction. Therefore, g_i and $\hat{B}g_i$ are linearly dependent functions, i.e., they are proportional,

$$\hat{B}g_i = b_i g_i \tag{3.38}$$

and consequently g_i is an eigenfunction of $\hat{B}$. We have just proved that the eigenfunctions of $\hat{A}$ are also eigenfunctions of $\hat{B}$ when the eigenvalues of $\hat{A}$ are nondegenerate.

Now suppose that the eigenfunctions of $\hat{A}$ are r-fold degenerate, that is, there are r eigenfunctions of $\hat{A}$ with the same eigenvalue. While any linear combination of these eigenfunctions is an eigenfunction of $\hat{A}$ with the same eigenvalue (this can be easily proved; try the twofold degenerate case, for example), it is not certain that the chosen eigenfunctions of $\hat{A}$ are also eigenfunctions of $\hat{B}$. Linear combinations of the eigenfunctions of $\hat{A}$ that are eigenfunctions of $\hat{B}$ diagonalize the matrix of $\hat{B}$, since this matrix diagonalization finds linear combinations of the basis functions that correspond to eigenfunctions of $\hat{B}$. A complete derivation of Theorem 3 can be found in *Merzbacher*, Further Reading, Sect. 10.4.

Theorems 2 and 3 allow us to conclude that the commutativity of two linear operators is a necessary and sufficient condition for having a common complete set of eigenfunctions. Hermitian operators represent observables. Hence, when two Hermitian operators $\hat{A}$ and $\hat{B}$ commute, they share a common complete set of eigenfunctions. Hence, the corresponding eigenvalues a_i and b_i are definite and sharp values that can be simultaneously assigned to observables A and B.

Theorem 4 *Two eigenfunctions of a Hermitian operator that correspond to different eigenvalues are orthogonal.*

Consider the Hermitian operator $\hat{A}$ and the eigenvalue equations corresponding to different eigenvalues,

$$\hat{A}f = af \quad \hat{A}g = bg \tag{3.39}$$

Substitution of (3.39) in (3.28) leads to

$$(b - a)\int f^* g \mathrm{d}\tau = 0 \tag{3.40}$$

We conclude that f and g are orthogonal, since $b \neq a$.

Theorem 5 *Two independent eigenfunctions of a Hermitian operator that correspond to the same eigenvalue (degenerate eigenvalue) may not be orthogonal, but they can always be replaced by linear combinations that become orthogonal.*

Consider two independent eigenfunctions of the Hermitian operator $\hat{A}$ that correspond to the same eigenvalue (degenerate eigenvalue),

$$\hat{A}f = sf \quad \hat{A}g = sg \tag{3.41}$$

If we now assume that f and g are not orthogonal, we can take f and $g + cf$ (note that $g + cf$ is also an eigenfunction of $\hat{A}$ with the same eigenvalue s) and choose c such that these two independent eigenfunctions become orthogonal, i.e.,

$$\int f^*(g + cf)\mathrm{d}\tau = 0 \quad \therefore c = -\frac{\int f^* g \mathrm{d}\tau}{\int f^* f \mathrm{d}\tau} \tag{3.42}$$

In conclusion, a Hermitian operator has an orthonormal set of eigenfunctions whose eigenvalues are real numbers. This set of functions can be used to expand any quadratically integrable and well-behaved function with the same boundary conditions of the eigenfunctions; i.e., it is a complete set of functions.

3.1.5 Dirac Notation

A notation introduced by Dirac, very useful in quantum mechanics, uses the following symbolism:

$$\int f^* \hat{A} g \mathrm{d}\tau = \langle f | \hat{A} | g \rangle \quad \int f^* g \mathrm{d}\tau = \langle f | g \rangle \tag{3.43}$$

where the integrals are over the full range of coordinates of the wave functions. We have

$$\left(\int f^* g \mathrm{d}\tau \right)^* = \int f g^* \mathrm{d}\tau \tag{3.44}$$

so

$$\langle f | g \rangle^* = \langle g | f \rangle \tag{3.45}$$

3.2 Harmonic Oscillator

The atoms of a molecule oscillate about their equilibrium positions. For an isolated molecule (a molecule in the gas phase at low pressure), suitable spectroscopic methods show that these oscillations have well-defined frequencies. In order to understand the vibrations of molecules, we begin by separating the translations and rotations. For a diatomic molecule AB, the motions of atoms A and B involve six

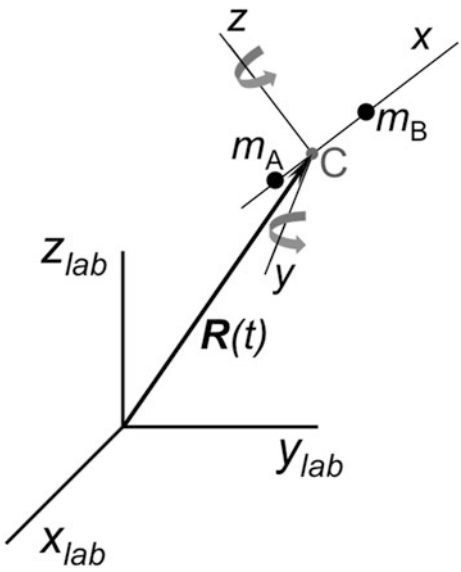


Fig. 3.1 In the diatomic molecule AB, the motions of atoms A and B can be described using two reference systems, one fixed in the laboratory (x_{lab}, y_{lab}, and z_{lab} axes), the other locked to the molecule with one of its axes coincident with the internuclear axis (C is the center of mass)

Cartesian coordinates, three for each atom. Let us consider two systems of reference, one fixed in the laboratory, the other locked to the molecule, with one of its axes coincident with the internuclear axis (Fig. 3.1). The motions of the atoms can now be separated into the center of mass motion described by three translations, corresponding to the components of the vector $\mathbf{R}(t)$ along the x_{lab}, y_{lab}, and z_{lab} axes, two rotations, around two orthogonal axes perpendicular to the internuclear axis (for a linear molecule, rotation around the internuclear axis does not lead to physically distinguishable configurations and so is not considered), and one vibration, described by the variation in the internuclear distance between atoms A and B (Fig. 3.1).

A molecule with N atoms needs $3N$ Cartesian coordinates as functions of time (three Cartesian coordinates per atom) to fully describe the motions of its atoms. Then, $3N$ coordinates include three coordinates for the center of mass motion (molecular translations), three coordinates for the rotations of a nonlinear molecule around three orthogonal axes, or two coordinates for the rotations of a linear molecule, and $3N - 6$ or $3N - 5$ coordinates for the vibrations of a nonlinear or a linear molecule, respectively.

3.2.1 Reduced Mass

Consider now a diatomic molecule AB as a system of two point masses defined by the vectors $\boldsymbol{r}_A$ and $\boldsymbol{r}_B$. The total kinetic energy is the sum of the kinetic energies of atoms A and B, that is,

$$T = \frac{1}{2} m_{\mathrm{A}} \left(\frac{\mathrm{d}\boldsymbol{r}_{\mathrm{A}}}{\mathrm{d}t} \right)^2 + \frac{1}{2} m_{\mathrm{B}} \left(\frac{\mathrm{d}\boldsymbol{r}_{\mathrm{B}}}{\mathrm{d}t} \right)^2 \tag{3.46}$$

If we now consider the molecule AB with its center of mass defined by the vector $\boldsymbol{R}$, then we can write

$$\boldsymbol{R}M = \boldsymbol{r}_{\mathrm{A}} m_{\mathrm{A}} + \boldsymbol{r}_{\mathrm{B}} m_{\mathrm{B}} \tag{3.47}$$

where

$$M = m_{\mathrm{A}} + m_{\mathrm{B}} \tag{3.48}$$

The stretching vibration is described by the following coordinate:

$$\boldsymbol{r} = \boldsymbol{r}_{\mathrm{B}} - \boldsymbol{r}_{\mathrm{A}} \tag{3.49}$$

Substitution of (3.47), (3.48), and (3.49) in (3.46) leads to

$$T = \frac{1}{2} M \left(\frac{d\boldsymbol{R}}{\mathrm{d}t} \right)^2 + \frac{1}{2} \mu \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \right)^2 \tag{3.50}$$

where μ is called the **reduced mass**, which is defined by the equality

$$\frac{1}{\mu} \equiv \frac{1}{m_{\mathrm{A}}} + \frac{1}{m_{\mathrm{B}}} \tag{3.51}$$

(**E12**, **E13**).

3.2.2 Classical Treatment

In the classical one-dimensional **harmonic oscillator** model, a single point mass m is connected to a rigid wall by a perfectly elastic spring (Fig. 3.2, left). The rigid wall is equivalent to an infinite mass. When the perfectly elastic spring is distended or contracted, the force that restores equilibrium is proportional to the shift from the equilibrium position,

$$F = -kx \tag{3.52}$$

where $x(t) = s(t) - s_{\mathrm{e}}$, with s_{e} representing the equilibrium position. The proportionality constant k is called a **force constant**. Use of the Newton equation in (3.52) gives

$$m \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -kx \tag{3.53}$$

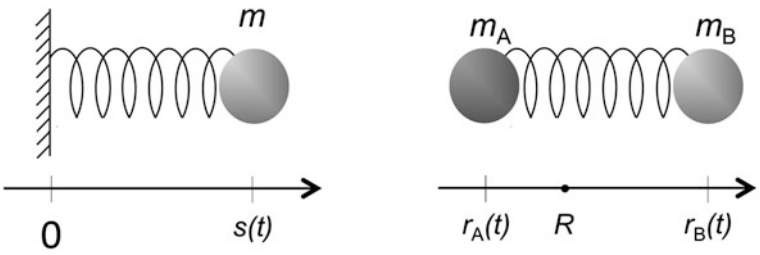


Fig. 3.2 (*Left*) A single mass linked to a rigid wall by a spring. (*Right*) Two masses linked by a spring (*R* is the center of mass of the system)

This is the differential equation for the classical one-dimensional harmonic oscillator. According to (3.53), the second derivative of $x(t)$ is proportional to $x(t)$, with the proportionality constant being equal to $-k/m$. Therefore, a sine or a cosine function or a linear combination of them satisfies (3.53). Note that the sine and cosine functions differ by a phase shift of $\pi/2$, and the following sine function,

$$x = A\sin(\omega t + b) \tag{3.54}$$

is equivalent to a linear combination of sine and cosine functions, where ω is the angular frequency of the oscillator ($\omega = 2\pi\nu$; units radian s^{-1}), while A and b are constants that result from integration of (3.53) (**E14**). The maximum and minimum values of x (A and $-A$) correspond to the maximum and minimum of the sine function in (3.54) (1 and −1). Substitution of (3.54) in (3.53) leads to

$$\omega = \sqrt{\frac{k}{m}} \tag{3.55}$$

3.2.3 Quantum-Mechanical Treatment

We now write

$$F = -\frac{\mathrm{d}V}{\mathrm{d}x} \tag{3.56}$$

where V is the potential energy. Substitution of (3.56) in (3.52) gives

$$\frac{\mathrm{d}V}{\mathrm{d}x} = kx \tag{3.57}$$

whose integration leads to

$$V = \frac{1}{2}kx^2 \tag{3.58}$$

on the assumption that the integration constant is zero (the potential energy always has an arbitrary additive constant, so we can take it as zero). The Taylor expansion of the potential energy around the potential energy minimum, $V_{\min}$, is given by

$$V = V_{\min} + \left(\frac{\mathrm{d}V}{\mathrm{d}x}\right)_{\min} + \frac{1}{2}\left(\frac{\mathrm{d}^2V}{\mathrm{d}x^2}\right)_{\min} x^2 + \cdots \tag{3.59}$$

The first term can be chosen to be zero, and the second term is zero because the minimum is a stationary point, and the slope at a stationary point is zero. Therefore, we can write, after taking (3.58) and (3.59) into account,

$$\left(\frac{\mathrm{d}^2V}{\mathrm{d}x^2}\right)_{\min} = k \tag{3.60}$$

that is, the curvature of the potential energy curve in the minimum is equal to the harmonic oscillator force constant. Terms of order higher than x^2 account for anharmonicity and are not included in the harmonic oscillator model.

The one-dimensional Schrödinger equation for the harmonic oscillator [(3.23) with the potential energy given by (3.58)] is

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_v(x)}{\mathrm{d}x^2} + \frac{1}{2}kx^2\psi_v(x) = E_v\psi_v(x) \tag{3.61}$$

where v is the vibrational quantum number (v = 0, 1, 2, …). After x is substituted by z using $x = z/\sqrt{\alpha}$ with $\alpha \equiv m\omega/\hbar$, the eigenfunctions for (3.61) become expressed as

$$\psi_v = N_v \mathrm{e}^{-z^2/2} H_v(z) \tag{3.62}$$

where the normalization constants are given by

$$N_v = \left(\frac{1}{\pi}\right)^{1/4}\left(\frac{1}{\sqrt{2^v v!}}\right) \tag{3.63}$$

(see *Mathematica* code for obtaining these normalization factors in **M3**), and H_v are the polynomial factors known as **Hermite polynomials** (**E15**, **E16**), defined by

$$H_v(z) \equiv (-1)^v \mathrm{e}^{z^2}\frac{\mathrm{d}^v}{\mathrm{d}z^v}\mathrm{e}^{-z^2} \tag{3.64}$$

These polynomials satisfy the recurrence formula

$$H_{v+1}(z) = 2zH_v(z) - 2vH_{v-1}(z) \tag{3.65}$$

(**E17**), and the first five are given by

$$\begin{aligned} H_0 &= 1 \\ H_1 &= 2z \\ H_2 &= 4z^2 - 2 \\ H_3 &= 8z^3 - 12z \\ H_4 &= 16z^4 - 48z^2 + 12 \end{aligned} \tag{3.66}$$

Figure 3.3 shows the squares of the harmonic oscillator eigenfunctions (3.62) (probability density curves), for v from 0 to 5 (see *Mathematica* code in **M3**). The number of nodes in each eigenfunction equals the value of the vibrational quantum number.

If we now substitute back z by x using $z = \sqrt{\alpha}x$, we obtain the first four harmonic oscillator normalized wave functions expressed in the original x variable:

$$\begin{aligned} \psi_0 &= \left(\frac{\alpha}{\pi}\right)^{1/4} \mathrm{e}^{-\alpha x^2/2} \\ \psi_1 &= \left(\frac{\alpha}{\pi}\right)^{1/4} \left(\frac{1}{\sqrt{2}}\right) \mathrm{e}^{-\alpha x^2/2} \left(2\sqrt{\alpha}x\right) \\ \psi_2 &= \left(\frac{\alpha}{\pi}\right)^{1/4} \left(\frac{1}{\sqrt{8}}\right) \mathrm{e}^{-\alpha x^2/2} \left(4\alpha x^2 - 2\right) \\ \psi_3 &= \left(\frac{\alpha}{\pi}\right)^{1/4} \left(\frac{1}{\sqrt{48}}\right) \mathrm{e}^{-\alpha x^2/2} \left[8\left(\sqrt{\alpha}\right)^3 x^3 - 12\sqrt{\alpha}x\right] \end{aligned} \tag{3.67}$$

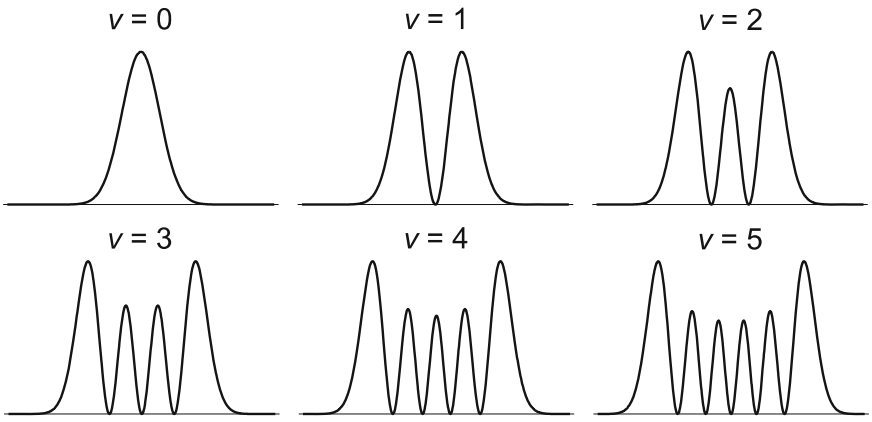


Fig. 3.3 Probability density curves for the harmonic oscillator as functions of z. This figure was obtained with *Mathematica*

where the normalized factors are $(\alpha/\pi)^{1/4}\,[1/\sqrt{(2^v v!)}]$ (see *Mathematica* code in **M3** for verifying these normalizing factors and confirming that these wave functions are orthogonal).

The existence of boundary conditions for the wave functions (these must go to zero as x goes to infinity) implies that $E_v = (v + 1/2)\hbar\omega$, where $v = 0, 1, 2, \ldots$, that is, the energies of the vibrational levels depend linearly on the vibrational quantum number v, and the separation between consecutive energy levels is $\hbar\omega$. At 0 K, the harmonic oscillator quantum number is zero ($v = 0$); the harmonic oscillator energy is called the **zero-point energy** and is equal to $\hbar\omega/2$. For a molecule as a collection of harmonic oscillators, the total zero-point energy is the sum of all zero-point energies.

3.2.4 Morse Potential

The single vibrational mode of a diatomic molecule is the stretching of its internuclear distance. An empirical representation of the potential energy function for a diatomic molecule, more realistic than the harmonic oscillator, is the so-called Morse potential energy function

$$U(R) = D\left[1 - \mathrm{e}^{-\beta(R-\mathrm{R_e})}\right]^2 \tag{3.68}$$

The Taylor expansion of this function about the equilibrium point R_e is

$$U(R) = U(R_\mathrm{e}) + (R - R_\mathrm{e})U'(R_\mathrm{e}) + \frac{(R - R_\mathrm{e})^2}{2!}U''(R_\mathrm{e}) + \frac{(R - R_\mathrm{e})^3}{3!}U'''(R_\mathrm{e}) + \cdots \tag{3.69}$$

where

$$U(R_\mathrm{e}) = 0 \quad \left(\frac{\mathrm{d}U}{\mathrm{d}R}\right)_{R_\mathrm{e}} = 0 \quad \left(\frac{\mathrm{d}^2U}{\mathrm{d}R^2}\right)_{R_\mathrm{e}} = k$$
$$\left(\frac{\mathrm{d}^3U}{\mathrm{d}R^3}\right)_{R_\mathrm{e}} = -3k\sqrt{\frac{k}{2D}} \quad \left(\frac{\mathrm{d}^4U}{\mathrm{d}R^4}\right)_{R_\mathrm{e}} = \frac{7k^2}{2D} \tag{3.70}$$

where $k = 2D\beta^2$ (**E18**). Substitution of (3.70) in (3.69) leads to

$$U(R) = \frac{k}{2}(R - R_\mathrm{e})^2 - \frac{k}{2}\sqrt{\frac{k}{2D}}(R - R_\mathrm{e})^3 + \frac{7k^2}{48D}(R - R_\mathrm{e})^4 + \cdots \tag{3.71}$$

The first nonzero term is the harmonic term. Cubic, quartic, and following terms account for the Morse potential anharmonicity.

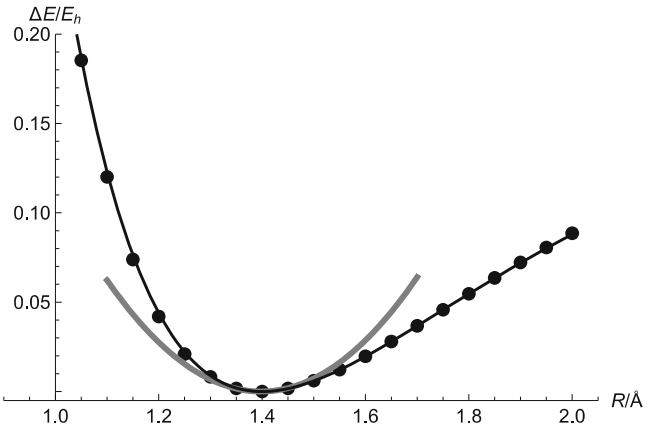


Fig. 3.4 Morse function fitted to the points obtained by scanning the bond distance of F_2. The parabola represents the harmonic oscillator potential energy curve. This figure was obtained with *Mathematica*

Figure 3.4 shows the potential energy curves for the harmonic oscillator (parabola) and the Morse function fitted to the points in black obtained by determining the potential energy of the system of two fluorine atoms at predefined internuclear distances (scan of the bond distance for F_2, Gaussian 09 B3LYP/cc-pVTZ calculation; for now, the student should simply consider this a reliable and accurate calculation for its purpose, leaving the details of the level of calculation and the basis of functions to be understood later on, in Sect. 3.15). The minimum of the parabola was set to be coincident with the minimum of the potential energy curve for F_2. Both curves (Morse function and parabola) have the same curvature at their minima. It can be seen that the harmonic oscillator deviates increasingly from the calculated potential energy curve for F_2 as the amplitude of vibration increases. Large-amplitude vibrations such as the stretching vibrations involving light atoms as in H_2, CH, NH, OH, FH, F_2, have important **anharmonic** corrections. Unlike the harmonic oscillator curve, the Morse potential function can be well fitted to the points of the potential energy curve of F_2.

3.3 Spherical Coordinates

The potential energy function for the electron of a hydrogen atom is a function of the distance from the electron at point (x, y, z) to the origin at the nucleus, and thus a spherically symmetric function. One says that the hydrogen atom is an example of a **central force** problem. The set of coordinates adequate to a central force problem are **spherical coordinates**, (r, θ, ϕ), where r is the distance from the origin to the point (x, y, z) $(0 \leq r < \infty)$, θ is the angle between the positive z-axis and the $\boldsymbol{r}$ vector to the point $(0 \leq \theta \leq \pi)$, and ϕ is the angle between the positive x-axis

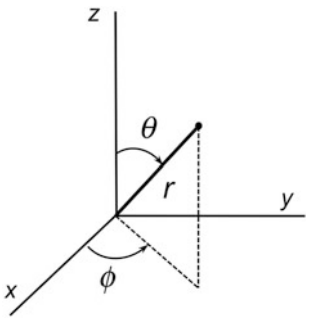


Fig. 3.5 Cartesian and spherical coordinates

with the projection of the $\boldsymbol{r}$ vector in the xy-plane ($0 \leq \phi < 2\pi$) (Fig. 3.5). Right triangle trigonometry can easily lead to the following equalities:

$$x = r\sin\theta\cos\phi \quad y = r\sin\theta\sin\phi \quad z = r\cos\theta \tag{3.72}$$

From these expressions, one can obtain

$$r = (x^2 + y^2 + z^2)^{1/2} \quad \cos\theta = \frac{z}{(x^2 + y^2 + z^2)^{1/2}} \quad \tan\phi = \frac{y}{x} \tag{3.73}$$

In Cartesian coordinates, the infinitesimal volume corresponds to a parallelepiped with edges equal to dx, dy, and dz, that is, dv = dxdydz. In spherical coordinates, two of the edges of the infinitesimal volume are curved, since they are associated with angular coordinates.

We begin by considering θ and ϕ constant. The intersection of the horizontal circle marked in Fig. 3.6 (r, θ constants) with the vertical semicircle that contains the z-axis and the projection of the $\boldsymbol{r}$ vector on the xy-plane (r, ϕ constants) is a line segment of length r (Fig. 3.6). The length of the infinitesimal edge between r and r + dr (not shown in Fig. 3.6) is given by dr,

$$\theta, \phi = \text{constant} \quad r \to r + \mathrm{d}r \quad \text{edge}_1 = \mathrm{d}r \tag{3.74}$$

We now consider r and ϕ constant. The intersection of the spherical surface (r constant) with the vertical plane that contains the z-axis and the projection of the $\boldsymbol{r}$ vector on the xy-plane (ϕ constant) is a vertical semicircle marked by a continuous line in Fig. 3.6. An infinitesimal arc between θ and θ + dθ (not shown in Fig. 3.6) is given by rdθ,

$$r, \phi = \text{constant} \quad \theta \to \theta + \mathrm{d}\theta \quad \text{edge}_2 = r\mathrm{d}\theta \tag{3.75}$$

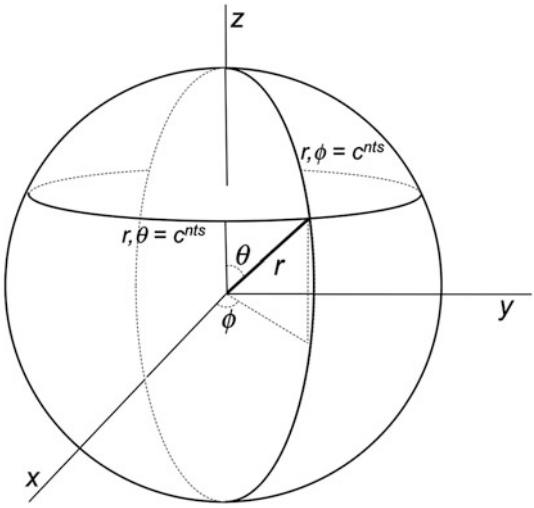


Fig. 3.6 The vertical semicircle corresponds to constant r and ϕ variables, whereas the horizontal circle corresponds to constant r and θ

The last edge of the infinitesimal volume in spherical coordinates is obtained by considering r and θ constant. The intersection of the spherical surface (r constant) with the conical surface (θ constant) is a horizontal circle marked in Fig. 3.6. An arc of this circle between ϕ and ϕ + dϕ (not shown in Fig. 3.6) is given by $r\sin\theta \mathrm{d}\phi$ (r sinθ is the projection of the $\boldsymbol{r}$ vector on the xy-plane),

$$r, \theta = \text{constant} \quad \phi \rightarrow \phi + \mathrm{d}\phi \quad \text{edge}_3 = r \sin\theta \mathrm{d}\phi \tag{3.76}$$

Thus, the infinitesimal volume in spherical coordinates is given by the product $\text{edge}_1 \times \text{edge}_2 \times \text{edge}_3$, which, after substitution of (3.74), (3.75), and (3.76), leads to

$$dv_{sph} = r^2 \sin\theta dr d\theta d\phi \tag{3.77}$$

3.4 Angular Momentum

In classical mechanics, the motion of one particle around a center of force can be characterized by two kinds of momenta, linear and angular momenta. Let $\boldsymbol{r}$ be the vector from the center of force to the particle position,

$$\boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z \tag{3.78}$$

and $\boldsymbol{v}$ the corresponding velocity vector,

$$\boldsymbol{v} \equiv \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{i}\frac{\mathrm{d}x}{\mathrm{d}t} + \boldsymbol{j}\frac{\mathrm{d}y}{\mathrm{d}t} + \boldsymbol{k}\frac{\mathrm{d}z}{\mathrm{d}t} \tag{3.79}$$

The classical definitions of the linear and angular momenta are given by

$$\boldsymbol{p} \equiv m\boldsymbol{v} \tag{3.80}$$

and

$$\boldsymbol{L} \equiv \boldsymbol{r} \times \boldsymbol{p} \tag{3.81}$$

i.e.,

$$\boldsymbol{L} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix} \tag{3.82}$$

where the components of $\boldsymbol{L}$ along the x, y, and z axes are

$$L_x = yp_z - zp_y \quad L_y = zp_x - xp_z \quad L_z = xp_y - yp_x \tag{3.83}$$

(**E19**).

3.4.1 Orbital Angular Momentum

The quantum-mechanical operators of the **linear momentum** [see (3.12)] are given by

$$\hat{p}_x = -i\hbar\frac{\partial}{\partial x} \quad \hat{p}_y = -i\hbar\frac{\partial}{\partial y} \quad \hat{p}_z = -i\hbar\frac{\partial}{\partial z} \tag{3.84}$$

In turn, the angular momentum associated with one-electron motion around a center of force is called **orbital angular momentum**. Substitution of (3.84) in (3.83) leads to the following expressions for the orbital angular momentum operators:

$$\hat{L}_x = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \quad \hat{L}_y = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) \quad \hat{L}_z = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \tag{3.85}$$

From these operators we can arrive at

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hbar\hat{L}_z \quad \left[\hat{L}_y, \hat{L}_z\right] = i\hbar\hat{L}_x \quad \left[\hat{L}_z, \hat{L}_x\right] = i\hbar\hat{L}_y \tag{3.86}$$

In order to derive these equalities, the above commutators should be applied to an arbitrary wave function f, and the equality between mixed partial second derivatives, $\partial^2 f/\partial x\partial y = \partial^2 f/\partial y\partial x$, should be used (**E20**). Note that (3.83), (3.85), and (3.86) obey the circular symmetry $x \to y \to z \to x$.

The operator for the square of angular momentum magnitude is given by

$$\hat{L}^2 = \hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{L}} = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \tag{3.87}$$

This operator commutes with each angular momentum component,

$$\left[\hat{L}^2, \hat{L}_x\right] = 0 \quad \left[\hat{L}^2, \hat{L}_y\right] = 0 \quad \left[\hat{L}^2, \hat{L}_z\right] = 0 \tag{3.88}$$

(**E21**). In conclusion, the square of the angular momentum operator of an isolated atom commutes with each of its components. However, no two components commute with each other. Therefore, the square of the angular momentum operator and only one of its angular momentum components, for example the z-component, have a common complete set of eigenfunctions. In spherical coordinates, the angular momentum operators involve only the variables θ and ϕ. In particular,

$$\hat{L}^2 = -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot\theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \phi^2} \right) \tag{3.89}$$

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi} \tag{3.90}$$

where the last expression suggests a separation of the variables θ and ϕ in the eigenfunctions. The one-electron eigenvalue equations and quantum numbers for the square and the z-component of the angular momentum operators are given by

$$\begin{aligned} &\hat{L}^2 Y_\ell^m(\theta,\phi) = \ell(\ell+1)\hbar^2 Y_\ell^m(\theta,\phi) \quad \ell = 0, 1, 2, \ldots \\ &\hat{L}_z Y_\ell^m(\theta,\phi) = m\hbar\, Y_\ell^m(\theta,\phi) \quad m = -\ell, -\ell+1, \ldots, 0, \ldots, \ell-1, \ell \end{aligned} \tag{3.91}$$

where the eigenfunctions

$$Y_\ell^m(\theta,\phi) = \sqrt{\frac{2\ell+1(\ell-m)!}{4\pi(\ell+m)!}} P_\ell^m(\cos\theta) e^{im\phi} \tag{3.92}$$

are called **spherical harmonics** and P_ℓ^m are the **associated Legendre polynomials** (note that in Y_ℓ^m and P_ℓ^m, m is a superscript, not an exponent) (see **§2**). Concise and useful definitions of associated Legendre polynomials, implemented in *Mathematica* as `LegendreP[`ℓ`,m, Cos[`θ`]]`, and spherical harmonics, implemented in *Mathematica* as `SphericalHarmonicY[`ℓ`,m,`θ`,`ϕ`]`, can be obtained by querying Wolfram Alpha at www.wolframalpha.com. The interested reader can find

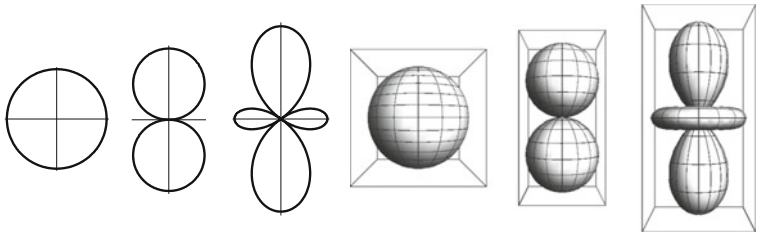


Fig. 3.7 Polar plots of associated Legendre polynomials P_0^0, P_1^0, and P_2^0, and three-dimensional plots of their corresponding spherical harmonics, obtained with *Mathematica*

in Wolfram MathWorld, at mathworld.wolfram.com, more detailed mathematical explanations on the associated Legendre polynomials and spherical harmonics. The *Mathematica* code **M4** presents the spherical harmonics expressions for $\ell \leq 2$ and the spherical harmonics 3D-plots for $\ell \leq 2$ and $m = 0$.

The obvious way of plotting a function $f(\theta)$ depicts θ as an angle and $|f(\theta)|$ as the distance from the origin to the point $[\theta, |f(\theta)|]$. Such a plot is called a **polar plot**. Since the distance from the point to the origin is the absolute value $|f(\theta)|$, the sign of $f(\theta)$ has to be separately marked over each region of space. Note that a polar plot emphasizes the *symmetry* of the plotted function, not its shape. For example, the polar plot for a constant is a circle, not a horizontal line. Figure 3.7 shows polar plots of the associated Legendre polynomials

$$P_0^0(\cos\theta) = 1 \quad P_1^0(\cos\theta) = \cos\theta \quad P_2^0(\cos\theta) = \frac{1}{2}(3\cos^2\theta - 1)$$

and three-dimensional plots of the corresponding spherical harmonics (**E22**, **E23**).

3.4.2 Spin

The yellow color imparted to a flame by sodium atoms results from emission from the electronic configuration $1s^2 2s^2 2p^6 3p^1$ to the ground state configuration $1s^2 2s^2 2p^6 3s^1$. While this spectral feature of atomic sodium atoms is usually called the sodium D line, it is in fact associated to two closely spaced spectral lines. In 1925, Uhlenbeck and Goudsmit explained this fine structure by proposing that the electron has an intrinsic angular momentum called the **spin angular momentum**, or more simply, **spin**, with two possible values of its projection along a preferred axis such as the axis of an applied magnetic field (Fig. 3.8). Later, in 1928, Dirac developed the relativistic quantum mechanics of the electron, where the electron spin arises naturally. In nonrelativistic quantum mechanics, the spin must be considered an additional hypothesis. While the orbital angular momentum we have

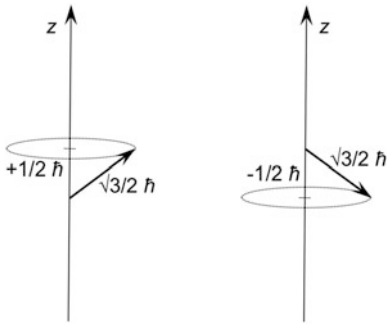


Fig. 3.8 Possible orientations of the spin vector lie on the surface of cones whose axes coincide with the z-axis

presented above is classically associated to one-electron motion around a center of force, the spin angular momentum has no classical mechanics equivalent.

The spin angular momentum operators are postulated to be Hermitian operators and obey commuting relationships analogous to the orbital angular momentum equivalents. For one electron, the spin eigenvalue equations and spin quantum numbers are

$$\begin{aligned} \hat{S}^2\psi &= s(s+1)\hbar^2\psi \quad s = 1/2 \\ \hat{S}_z\psi &= m_s\hbar\psi \quad m_s = 1/2, -1/2 \end{aligned} \tag{3.93}$$

where the +1/2 and −1/2 values of m_s correspond to α and β spin eigenfunctions that form a complete set of orthonormal functions, i.e., normalized and mutually orthogonal functions,

$$\int \alpha^* \alpha \mathrm{d}\sigma = \int \beta^* \beta \mathrm{d}\sigma = 1 \quad \int \alpha^* \beta \mathrm{d}\sigma = \int \beta^* \alpha \mathrm{d}\sigma = 0 \tag{3.94}$$

Since the spin is a quantum concept without classical equivalent, (3.94) should be seen as merely expressing formal equations.

3.5 Hydrogen Atom

The Hamiltonian for the hydrogen atom is given by

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla^2 + V(r) \tag{3.95}$$

where

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{3.96}$$

is the Laplacian operator and μ is the reduced mass of the electron–nucleus system that satisfies the following equality:

$$\frac{1}{\mu} \equiv \frac{1}{m_\mathrm{e}} + \frac{1}{m_\mathrm{N}} \tag{3.97}$$

For the lightest atom protium, the mass of the nucleus is about 2×10^3 times the mass of the electron. Therefore, the second term on the right-hand side of (3.97) is negligible, and μ is approximately equal to the electron mass,

$$\mu \approx m_\mathrm{e} \tag{3.98}$$

The potential energy operator $V(r)$ is given by

$$V(r) = -\frac{\mathrm{e}^2}{4\pi\varepsilon_0 r} \tag{3.99}$$

where r is the electron–nucleus distance and ε_0 is the vacuum permittivity (see Appendix).

In spherical coordinates, the Laplacian operator takes the form

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{1}{r^2\hbar^2}\hat{L}^2 \tag{3.100}$$

Therefore, the Hamiltonian for the hydrogen atom commutes with the square of the angular momentum operator [see (3.89)] and with the z-component of the angular momentum operator [see (3.90)], and we can write

$$[\hat{H}, \hat{L}^2] = 0 \quad [\hat{H}, \hat{L}_z] = 0 \tag{3.101}$$

(**E24**). These equalities allow us to conclude that there is a complete set of eigenfunctions common to the Hamiltonian for the hydrogen atom, the square of the angular momentum operator, and its z-component. The square of the angular momentum operator and its z-component do not involve the coordinate r [see (3.89) and (3.90)], so the radial factors act like constants with respect to the angular momentum operators. Hence, the eigenfunctions for the hydrogen atom are obtained by multiplying radial functions by spherical harmonics,

$$\psi_{n\ell m} = R_{n\ell}(r) Y_\ell^m(\theta, \phi) \tag{3.102}$$

with the radial factors and spherical harmonics being labeled by the n, ℓ, and m quantum numbers. Alternatively, ℓ can be substituted by a letter, where the first letters of the words of spectroscopic origin sharp, principal, diffuse, and fundamental stand for $\ell = 0$, 1, 2, and 3, respectively. After the letter f, the alphabetical order is followed, but j is omitted, and s and p are skipped because they have already been used for $\ell = 0$ and 1,

$$\begin{array}{cccccccc} \ell & 0 & 1 & 2 & 3 & 4 & \dots \\ & s & p & d & f & g & \dots \end{array}$$

Boundary conditions for the radial factors imply the introduction of the quantum number n. For $n \leq 3$ and the hydrogen atom ($Z = 1$) wave functions in atomic units (unit of length is $a_0 = 0.52917721092 \times 10^{-10}$ m; see Appendix), the radial factors are given by

$$\begin{aligned} R_{1s} &= 2\mathrm{e}^{-r} \\ R_{2s} &= \frac{1}{2\sqrt{2}}(2-r)\mathrm{e}^{-r/2} \quad R_{2p} = \frac{1}{2\sqrt{6}} r\mathrm{e}^{-r/2} \\ R_{3s} &= \frac{2}{81\sqrt{3}}(27-18r+2r^2)\mathrm{e}^{-r/3} \quad R_{3p} = \frac{4}{81\sqrt{6}}(6-r)r\mathrm{e}^{-r/3} \quad R_{3d} = \frac{4}{81\sqrt{30}} r^2\mathrm{e}^{-r/3} \end{aligned} \tag{3.103}$$

(**E25, E26**).

The probability of finding the electron with spherical coordinates lying in the ranges r to $r + \mathrm{d}r$, θ to $\theta + \mathrm{d}\theta$, and ϕ to $\phi + \mathrm{d}\phi$ is given by

$$|\psi_{n\ell m}(r,\theta,\phi)|^2 r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi = R_{n\ell}^2(r) r^2 \mathrm{d}r \left|Y_\ell^m(\theta,\phi)\right|^2 \sin\theta\mathrm{d}\theta\mathrm{d}\phi \tag{3.104}$$

The probability of finding the electron in the range r to $r + \mathrm{d}r$, independently of θ and ϕ, is given by

$$R_{n\ell}^2(r) r^2 \mathrm{d}r \tag{3.105}$$

because the spherical harmonics are normalized functions, that is,

$$\int_0^{2\pi} \int_0^{\pi} \left|Y_\ell^m(\theta,\phi)\right|^2 \sin\theta\mathrm{d}\theta\mathrm{d}\phi = 1 \tag{3.106}$$

The $|R_{n\ell}(r)|^2 r^2$ are called **radial distribution** functions (Fig. 3.9).

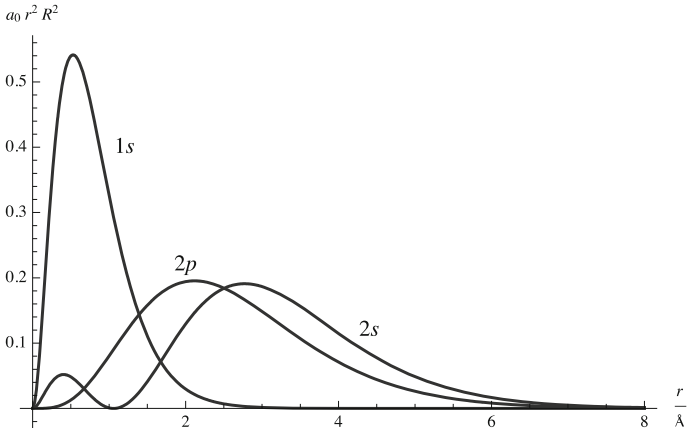


Fig. 3.9 Radial distribution functions for the 1*s*, 2*s*, and 2*p* hydrogen orbitals. Compare 2*s* and 2*p* and note the 2*s* lobes separated by a node. This graph was obtained with *Mathematica*

3.6 Antisymmetry Principle

In classical mechanics, each point of a particle trajectory is described by its position and linear momentum coordinates (six coordinates). Identical particles can be distinguished by distinct trajectories. However, electrons cannot be assigned individual trajectories due to the uncertainty principle that states that position and momentum cannot be simultaneously specified. Each electron of an atom or a molecule is assigned four variables: three spatial coordinates that specify the position vector (for example, x, y, z) and the spin variable that takes the value α for $m_s = +1/2$ or β for $m_s = -1/2$. If we represent the set of these four variables by s, the wave function for two electrons can be written as $\Psi(s_1, s_2)$. Consider now the exchange or permutation operator that interchanges all four coordinates of electrons 1 and 2,

$$\hat{P}_{12}\Psi(s_1, s_2) = \Psi(s_2, s_1) \tag{3.107}$$

Applying this operator twice has no net effect, since the initial configuration is restored,

$$\hat{P}_{12}^2\Psi(s_1, s_2) = \hat{P}_{12}\left[\hat{P}_{12}\Psi(s_1, s_2)\right] = \hat{P}_{12}\Psi(s_2, s_1) = \Psi(s_1, s_2) \tag{3.108}$$

Therefore, the eigenvalue of the square of the permutation operator is 1, and consequently, the eigenvalue of the permutation operator is either +1 or −1. However, only −1 is consistent with experimental data for electrons, that is,

$$\hat{P}_{12}\Psi(s_1, s_2) = -\Psi(s_1, s_2) \tag{3.109}$$

Combining (3.107) and (3.109) leads to

$$\Psi(s_2, s_1) = -\Psi(s_1, s_2) \tag{3.110}$$

and the wave function is said to be **antisymmetric** with respect to the interchange of coordinates of two electrons including the spin coordinates (**antisymmetry principle**). If $s_2 = s_1$, then

$$\Psi(s_1, s_1) = -\Psi(s_1, s_1) \tag{3.111}$$

and consequently,

$$\Psi(s_1, s_1) = 0 \quad \therefore |\Psi(s_1, s_1)|^2 = 0 \tag{3.112}$$

Therefore, two electrons with the same spin have zero probability of being found at the same point. Since the wave function is a well-behaved function and so is continuous, the probability of finding two electrons with the same spin in the same region of space should be small. This consequence of the antisymmetry principle is called the **Pauli exclusion principle**.

How to build a many-electron wave function Ψ from one-electron functions, so that Ψ obeys the antisymmetry requirement? First of all, we take each one-electron function as being obtained by the product of a spatial factor and one of the spin functions α or β. For electron 1, we can write

$$\psi_a(s_1) = \phi_a(\boldsymbol{r}_1)\sigma(1) \tag{3.113}$$

where σ can be either α or β. Each one-electron spatial factor is called an **orbital**, and the product of a spatial factor and the corresponding spin factor is called a **spin-orbital**. In calculations, the same spatial factor is usually combined with different spin factors α and β.

Consider now a three-electron system and the product of three spin-orbitals,

$$\psi_a(1)\psi_b(2)\psi_c(3) \tag{3.114}$$

In order to make electrons indistinguishable, all permutations of the electron labels must be considered. In addition, for each permutation of the coordinates of two electrons, the sign must be changed, and so we obtain

$$\begin{aligned}&\psi_a(1)\psi_b(2)\psi_c(3) - \psi_a(2)\psi_b(1)\psi_c(3) + \psi_a(2)\psi_b(3)\psi_c(1)\\ &\quad - \psi_a(3)\psi_b(2)\psi_c(1) + \psi_a(3)\psi_b(1)\psi_c(2) - \psi_a(1)\psi_b(3)\psi_c(2)\end{aligned} \tag{3.115}$$

The one-electron functions ψ_i are orthonormal, so integration of the square of the absolute value of (3.115) gives 1 + 1 + 1 + 1 + 1 + 1 = 6, and the normalizing constant is $1/\sqrt{6} = 1/\sqrt{3!}$. Changing the simple product wave function (3.114)

into a normalized antisymmetric wave function Ψ can be formally represented by the operator $\mathcal{A}$ with the normalization constant included, where

$$\Psi = \mathcal{A}[\psi_a(1)\psi_b(2)\psi_c(3)] \tag{3.116}$$

Hence, applying $\mathcal{A}$ to a simple product of n one-electron orthonormal functions gives an antisymmetric linear combination of $n!$ permutations with the normalization constant $1/\sqrt{n!}$. The resulting normalized function Ψ is **antisymmetric** with respect to the exchange of a pair of electron coordinates, thus satisfying the antisymmetry requirement.

An efficient way of obtaining all permutations of the electron coordinates consists in taking the normalized "determinant" of one-electron functions called a **Slater determinant**. A 3×3 Slater determinant is given by

$$\Psi = \left(\frac{1}{\sqrt{3!}}\right) \begin{vmatrix} \psi_a(1) & \psi_b(1) & \psi_c(1) \\ \psi_a(2) & \psi_b(2) & \psi_c(2) \\ \psi_a(3) & \psi_b(3) & \psi_c(3) \end{vmatrix} \tag{3.117}$$

Permuting a pair of electron coordinates is equivalent to exchanging lines in the Slater determinant (each line of the Slater determinant corresponds to a particular electron coordinate). For instance, permuting rows 1 and 2 leads to

$$\begin{vmatrix} \psi_a(2) & \psi_b(2) & \psi_c(2) \\ \psi_a(1) & \psi_b(1) & \psi_c(1) \\ \psi_a(3) & \psi_b(3) & \psi_c(3) \end{vmatrix} = - \begin{vmatrix} \psi_a(1) & \psi_b(1) & \psi_c(1) \\ \psi_a(2) & \psi_b(2) & \psi_c(2) \\ \psi_a(3) & \psi_b(3) & \psi_c(3) \end{vmatrix} \tag{3.118}$$

[note that normalizing constants have been omitted in (3.118)]. In addition, two equal columns make the Slater determinant equal to zero, since

$$\begin{vmatrix} \psi_a(1) & \psi_a(1) & \psi_c(1) \\ \psi_a(2) & \psi_a(2) & \psi_c(2) \\ \psi_a(3) & \psi_a(3) & \psi_c(3) \end{vmatrix} = 0 \tag{3.119}$$

and the corresponding probability density becomes zero. Therefore, no two electrons of an atom or a molecule can have the same spin-orbital. This conclusion shows that the Pauli exclusion principle is a corollary of the antisymmetric behavior of the wave function.

If columns 1 and 3 of the Slater determinant in (3.117) are substituted by normalized orthogonal combinations of these columns, the resulting determinant remains the same or simply changes its sign (the corresponding probability density is not altered),

$$\begin{vmatrix} \frac{\psi_a(1)+\psi_c(1)}{\sqrt{2}} & \psi_b(1) & \frac{\psi_a(1)-\psi_c(1)}{\sqrt{2}} \\ \frac{\psi_a(2)+\psi_c(2)}{\sqrt{2}} & \psi_b(2) & \frac{\psi_a(2)-\psi_c(2)}{\sqrt{2}} \\ \frac{\psi_a(3)+\psi_c(3)}{\sqrt{2}} & \psi_b(3) & \frac{\psi_a(3)-\psi_c(3)}{\sqrt{2}} \end{vmatrix} = -\begin{vmatrix} \psi_a(1) & \psi_b(1) & \psi_c(1) \\ \psi_a(2) & \psi_b(2) & \psi_c(2) \\ \psi_a(3) & \psi_b(3) & \psi_c(3) \end{vmatrix} \tag{3.120}$$

The *Mathematica* code **M5** illustrates the above properties of determinantal functions and confirms (3.118), (3.119), and (3.120) for 3 × 3 matrices and their determinants.

Consider now the following Slater determinant for a two-electron atom or molecule:

$$\Psi(s_1, s_2) = \frac{1}{\sqrt{2!}} \begin{vmatrix} \psi_a(s_1) & \psi_b(s_1) \\ \psi_a(s_2) & \psi_b(s_2) \end{vmatrix} = \frac{1}{\sqrt{2}} [\psi_a(s_1)\psi_b(s_2) - \psi_a(s_2)\psi_b(s_1)] \tag{3.121}$$

Taking each spin-orbital as the product of a common spatial factor ϕ_a by α and β, one obtains a closed shell configuration and the following wave function:

$$\begin{aligned} \Psi(s_1, s_2)^{(S=0, M_S=0)} &= \frac{1}{\sqrt{2!}} \begin{vmatrix} \phi_a(\boldsymbol{r}_1)\alpha(1) & \phi_a(\boldsymbol{r}_1)\beta(1) \\ \phi_a(\boldsymbol{r}_2)\alpha(2) & \phi_a(\boldsymbol{r}_2)\beta(2) \end{vmatrix} \\ &= \phi_a(\boldsymbol{r}_1)\phi_a(\boldsymbol{r}_2) \frac{1}{\sqrt{2}} [\alpha(1)\beta(2) - \alpha(2)\beta(1)] \end{aligned} \tag{3.122}$$

Electrons 1 and 2 have antiparallel spins, so the sum $M_s = m_{s1} + m_{s2}$ is zero. Therefore, the total spin is zero ($S = 0$), and the number of distinct values of M_S, or **electron-spin multiplicity** $(2S + 1)$, is equal to 1 (the state is called a **singlet**).

Consider now an open shell configuration with electrons 1 and 2 in different orbitals, both with spins α. The corresponding Slater determinant is the product of an antisymmetric product of spatial factors and the simple product of spin functions:

$$\begin{aligned} \Psi(s_1, s_2)^{(S=1, M_S=1)} &= \frac{1}{\sqrt{2!}} \begin{vmatrix} \phi_a(\boldsymbol{r}_1)\alpha(1) & \phi_b(\boldsymbol{r}_1)\alpha(1) \\ \phi_a(\boldsymbol{r}_2)\alpha(2) & \phi_b(\boldsymbol{r}_2)\alpha(2) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}} [\phi_a(\boldsymbol{r}_1)\phi_b(\boldsymbol{r}_2) - \phi_a(\boldsymbol{r}_2)\phi_b(\boldsymbol{r}_1)]\alpha(1)\alpha(2) \end{aligned} \tag{3.123}$$

This electronic state corresponds to $M_s = m_{s1} + m_{s2} = 1/2 + 1/2 = 1$. If both spin factors are given by β, then $M_s = m_{s1} + m_{s2} = (-1/2) + (-1/2) = -1$, and the total wave function is

$$\begin{aligned} \Psi(s_1, s_2)^{(S=1, M_S=-1)} &= \frac{1}{\sqrt{2!}} \begin{vmatrix} \phi_a(\boldsymbol{r}_1)\beta(1) & \phi_b(\boldsymbol{r}_1)\beta(1) \\ \phi_a(\boldsymbol{r}_2)\beta(2) & \phi_b(\boldsymbol{r}_2)\beta(2) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}} [\phi_a(\boldsymbol{r}_1)\phi_b(\boldsymbol{r}_2) - \phi_a(\boldsymbol{r}_2)\phi_b(\boldsymbol{r}_1)]\beta(1)\beta(2) \end{aligned} \tag{3.124}$$

The wave function that corresponds to $M_s = m_{s1} + m_{s2} = 0$ multiplies an antisymmetric product of spatial factors by a linear combination of symmetric products of spin functions and can be expressed by the following linear combination of Slater determinants:

$$\begin{aligned}\Psi_{(S_1,S_2)}^{(S=1,M_S=0)} &= \frac{1}{2}\left\{\begin{vmatrix}\phi_a(\boldsymbol{r}_1)\alpha(1) & \phi_b(\boldsymbol{r}_1)\alpha(1)\\ \phi_a(\boldsymbol{r}_2)\beta(2) & \phi_b(\boldsymbol{r}_2)\beta(2)\end{vmatrix} + \begin{vmatrix}\phi_a(\boldsymbol{r}_1)\beta(1) & \phi_b(\boldsymbol{r}_1)\beta(1)\\ \phi_a(\boldsymbol{r}_2)\alpha(2) & \phi_b(\boldsymbol{r}_2)\alpha(2)\end{vmatrix}\right\}\\ &= \frac{1}{\sqrt{2}}[\phi_a(\boldsymbol{r}_1)\phi_b(\boldsymbol{r}_2) - \phi_a(\boldsymbol{r}_2)\phi_b(\boldsymbol{r}_1)]\frac{1}{\sqrt{2}}[\alpha(1)\beta(2) + \beta(1)\alpha(2)]\end{aligned} \tag{3.125}$$

Note that each individual determinant of (3.125) is not an eigenfunction of the square of the spin angular momentum operator and so does not correspond to a pure spin state. However, the linear combination of determinants (3.125) is an eigenfunction of the square of the spin angular momentum operator and is called a **spin-adapted** wave function. Wave functions (3.123), (3.124), and (3.125) belong to the same **triplet** ($S = 1$, $2S + 1 = 3$). For more than two electrons, the wave function cannot be factorized into spatial and spin coordinates.

3.7 Variational Method

Solving the Schrödinger equation for atoms or molecules with more than one electron requires the use of approximate methods, the most powerful of these being based on the variational theorem. Consider a system with a Hamiltonian whose lowest-energy eigenvalue is E_0, and a well-behaved function ϕ that satisfies the boundary conditions of the system. According to the variational theorem,

$$\frac{\int \phi^* \hat{H}\phi \mathrm{d}\tau}{\int \phi^* \phi \mathrm{d}\tau} \geq E_0 \tag{3.126}$$

The eigenfunctions ψ_k of the Hamiltonian satisfy the eigenvalue equations

$$\hat{H}\psi_k = E_k\psi_k \tag{3.127}$$

and form a complete orthonormal set of functions that can be used to expand ϕ,

$$\phi = \sum_k a_k \psi_k \tag{3.128}$$

To prove (3.126), we substitute (3.128) into (3.126) and use (3.127) to obtain

$$W \equiv \frac{\int \phi^* \hat{H}\phi \mathrm{d}\tau}{\int \phi^* \phi \mathrm{d}\tau} = \frac{\sum_j \sum_k a_j^* a_k E_k \delta_{jk}}{\sum_j \sum_k a_j^* a_k \delta_{jk}} = \frac{\sum_k a_k^* a_k E_k}{\sum_k a_k^* a_k} \geq E_0 \tag{3.129}$$

where δ_{jk} is the Kronecker delta ($\delta_{jk} = 0$, for $j \neq k$, and $\delta_{jk} = 1$, for $j = k$) and the inequality is justified because E_0 is the lowest-energy eigenvalue.

If ϕ is a linear combination of n linearly independent functions that satisfy the boundary conditions of the system,

$$\phi = \sum_{j=1}^{n} c_j f_j \tag{3.130}$$

the set of functions f_j is said to be **a set of basis functions** with the coefficients being the variation parameters. Assuming real basis functions and real coefficients, then

$$W = \frac{\int \phi^* \hat{H} \phi \mathrm{d}\tau}{\int \phi^* \phi \mathrm{d}\tau} = \frac{\sum_j \sum_k c_j c_k H_{jk}}{\sum_j \sum_k c_j c_k S_{jk}} \tag{3.131}$$

where

$$S_{jk} = \int f_j f_k \mathrm{d}\tau \quad H_{jk} = \int f_j \hat{H} f_k \mathrm{d}\tau \tag{3.132}$$

Note that W is a function of another function, in this case ϕ. Thus, W is called a **functional** of ϕ, and the functional dependence is represented by a square-brackets notation, $W[\phi]$. While a function is a rule that associates a number with another number [for example, $f(x) = 2x^2 + 3$ associates the number $x = 3$ with the number $2x^2 + 3 = 21$], a functional is a rule that associates a number (the functional value) with a function.

Minimization of $W[\phi]$ with respect to the coefficients can be achieved by solving the system of equations

$$\frac{\partial W}{\partial c_j} = 0 \tag{3.133}$$

which is equivalent to

$$\frac{\partial \Lambda}{\partial c_j} - W \frac{\partial \Delta}{\partial c_j} = 0 \tag{3.134}$$

where

$$\Lambda = \sum_j \sum_k c_j c_k H_{jk} \quad \Delta = \sum_j \sum_k c_j c_k S_{jk} \tag{3.135}$$

(**E27**). Illustration of the above equations with a set of two basis functions leads to

$$\begin{array}{ll} \frac{\partial \Lambda}{\partial c_1} = 2c_1H_{11} + 2c_2H_{12} & \frac{\partial \Lambda}{\partial c_1} = 2c_1S_{11} + 2c_2S_{12} \\ \frac{\partial \Lambda}{\partial c_2} = 2c_1H_{12} + 2c_2H_{22} & \frac{\partial \Lambda}{\partial c_2} = 2c_1S_{12} + 2c_2S_{22} \end{array} \tag{3.136}$$

The Hamiltonian is Hermitian, and we have assumed real basis functions, so $H_{12} = H_{21}$ and $S_{12} = S_{21}$. Substitution of (3.136) in (3.134) gives

$$\begin{cases} (H_{11} - WS_{11})c_1 + (H_{12} - WS_{12})c_2 = 0 \\ (H_{12} - WS_{12})c_1 + (H_{22} - WS_{22})c_2 = 0 \end{cases} \tag{3.137}$$

(**E28**, **E29**, **E30**). A nontrivial solution for this system of equations (the trivial solution $c_1 = c_2 = 0$ corresponds to a null ϕ everywhere) requires a vanishing determinant of the coefficients, i.e.,

$$\begin{vmatrix} H_{11} - WS_{11} & H_{12} - WS_{12} \\ H_{12} - WS_{12} & H_{22} - WS_{22} \end{vmatrix} = 0 \tag{3.138}$$

Expansion of (3.138) gives an equation of second degree in W whose roots are $W^{(1)}$ and $W^{(2)}$. The lowest root $W^{(1)}$ is an approximation to the lowest energy eigenvalue. Substitution of $W^{(1)}$ in (3.137) gives a system of linear homogeneous equations with the coefficients $c_1^{(1)}$ and $c_2^{(1)}$ as unknowns. The determinant of the coefficients is zero, so the equations are equivalent, that is, they differ by a constant factor. Therefore, we have two unknowns and only one independent equation. This system of equations is called underdetermined, and its solution gives the ratio of the coefficients, $c_2^{(1)}/c_1^{(1)}$. Normalization of $\phi^{(1)}$ enables one to determine $c_1^{(1)}$ and $c_2^{(1)}$.

A nontrivial solution for a system of n linear homogeneous equations in n unknowns requires a zero determinant corresponding to (3.138) in the 2×2 case. Being a system of homogeneous equations, the unknowns can be multiplied by an arbitrary constant. This means that we are left with n equations and $n - 1$ unknowns, which are the ratios c_2/c_1, c_3/c_1, …, c_n/c_1 of the coefficients. In order to solve for the values of all coefficients, we need to add a new equation that normalizes ϕ. The *Mathematica* code **M6** illustrates this process of solving systems of homogeneous linear equations with 2×2 and 3×3 coefficient matrices.

3.8 Born–Oppenheimer Approximation

Consider the nonrelativistic Hamiltonian for a molecule

$$\hat{H} = \underbrace{-\frac{\hbar^2}{2}\sum_{\alpha}\frac{\nabla_\alpha^2}{m_\alpha}}_{\hat{T}_N} \underbrace{-\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2}_{\hat{T}_e} + \frac{e^2}{4\pi\varepsilon_0}\left(\underbrace{\sum_{\alpha}\sum_{\beta>\alpha}\frac{Z_\alpha Z_\beta}{r_{\alpha\beta}}}_{\hat{V}_{NN}} \underbrace{-\sum_i\sum_{\alpha}\frac{Z_\alpha}{r_{\alpha i}}}_{\hat{V}_{Ne}} + \underbrace{\sum_i\sum_{j>i}\frac{1}{r_{ij}}}_{\hat{V}_{ee}} \right) \tag{3.139}$$

where α and β refer to nuclei, i and j to electrons. The first and second terms on the right side of (3.139) are operators for the kinetic energies of the nuclei and electrons, and the last three terms are potential energy operators for the nucleus–nucleus repulsions, electron–nucleus attractions, and electron–electron repulsions. Benzene, a typical medium-size molecule, has 12 nuclei and 42 electrons corresponding to 36 nuclear coordinates and 126 electron coordinates. Solving the Schrödinger equation for this molecule is a formidable task. However, nuclei are much heavier than electrons: one proton mass is about 1.8×10^3 times the electron mass, and one ^{12}C nucleus mass is about 2.2×10^4 times the electron mass. This suggests an approximation whereby electron and nuclear coordinates are dealt with separately, with the electrons carrying out their motions while the nuclei are fixed. The **Born–Oppenheimer approximation** deals with the separation between electron and nuclear coordinates in two steps. In the first step, the kinetic energy of the nuclei is assumed to be zero and the Schrödinger equation

$$(\hat{H}_e + \hat{V}_{NN})\psi_e(s_i, R_\alpha) = U(R_\alpha)\psi_e(s_i, R_\alpha) \tag{3.140}$$

is solved, where

$$\hat{H}_e = \hat{T}_e + \hat{V}_{Ne} + \hat{V}_{ee} \tag{3.141}$$

is the **electronic Hamiltonian**, s_i represents the electron coordinates (spins included), and R_α stands for the nuclear position coordinates. Multiplying both members of (3.140) by the complex conjugate of the wave function and integrating with respect to the electron coordinates leads to the potential energy $U(R_\alpha)$, which depends parametrically on the nuclear coordinates, $U(R_\alpha) = E_e + V_{NN}(R_\alpha)$, thus becoming the potential energy for the nuclear motion. The second step of the Born–Oppenheimer approximation considers the Schrödinger equation for the nuclear motion

$$\hat{H}_N\psi_N(R_\alpha) = E_N\psi_N(R_\alpha) \tag{3.142}$$

where

$$\hat{H}_N = \hat{T}_N + U(R_\alpha) \tag{3.143}$$

When (3.142) is solved using a molecule-fixed reference frame, the molecular rotations and translations are not accounted for, and the energy eigenvalues E_N correspond to internal degrees of freedom, which are the molecular vibrations.

For a diatomic molecule,

$$\hat{V}_{NN} = \frac{ZZ'e^2}{4\pi\varepsilon_0 R} \tag{3.144}$$

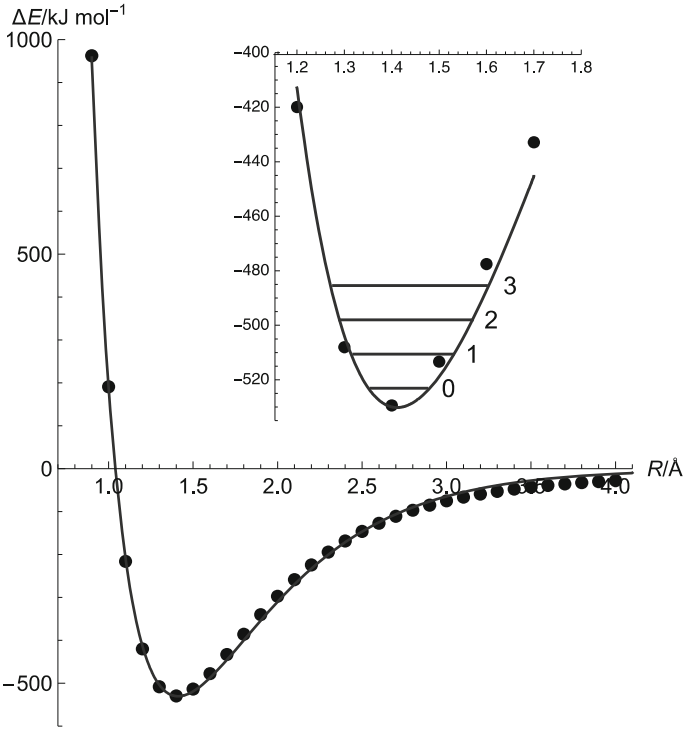


Fig. 3.10 A $U(R)$ function for the ground electronic state of F_2 obtained by scanning the internuclear distance (Gaussian 09 B3LYP/cc-pVTZ calculation). The inset shows the zero-point energy level marked with 0 and the first, second, and third stretching energy levels marked with 1, 2, and 3, respectively. This figure was obtained with *Mathematica*

where Z and Z' are atomic numbers and R is the internuclear distance. The potential energy curve $U(R)$ is a function of R, and the difference between the energy at infinite internuclear distance and the energy at the minimum R_e is called the **equilibrium dissociation energy**,

$$D_e = U(\infty) - U(R_e) \tag{3.145}$$

Figure 3.10 shows points obtained by solving the electronic Schrödinger equation (3.140) for two fluorine atoms at defined values of the internuclear distance R (scan of R). These points were obtained by scanning the internuclear distance R, and a curve was fitted to the points using a Morse potential to which R^{-6}, R^{-8}, and R^{-10} terms were added. The inset presents vibrational levels obtained by solving the nuclear Schrödinger equation (3.142), that is, by performing a frequency calculation. The energy level 0 corresponds to the zero-point energy, 0.002392 E_h = 6.28 kJ mol^{-1}, marked from the minimum of the curve (see the appendix for conversion factors). The 1–0 transition corresponds to the harmonic frequency 1049.88 cm^{-1} = 12.56 kJ mol^{-1}. Levels 2 and 3 were drawn assuming

harmonic behavior, that is, the 2–1 and 3–2 transitions were made equal to the harmonic frequency (anharmonicity was ignored). All these calculations were carried out using the system of programs *Gaussian* 09, with the B3LYP method of calculation and the cc-pVTZ basis set (the student can leave the technical details of this calculation to Sect. 3.15).

3.9 Hartree–Fock Method

The electron–electron repulsion operator of an atomic or molecular Hamiltonian includes the distance r_{ij} between any pair of electrons [see (3.139)]. This variable makes the electron coordinates interdependent, since each electron position depends on the positions of all other electrons due to instantaneous repulsions. The interdependence of the electron coordinates implies that the coordinates of the electrons cannot be separated, and consequently, one-electron wave functions (orbitals) are approximations, not exact functions. Hartree (1897–1958) developed a method to optimize one-electron functions as factors of many-electron simple product wave functions. Later, Slater (1900–1976) and Fock (1898–1974) extended the method to antisymmetric linear combinations of simple products (Slater determinants). In the Hartree and Hartree–Fock methods, instantaneous electron–electron repulsions are replaced by interactions between each electron and the average potential due to all other electrons. This approximation enables the use of one-electron functions (orbitals) in the formation of many-electron wave functions for atoms and molecules.

Consider an atom with two electrons and the following simple product of spatial orbitals:

$$\psi^{(0)}(1,2) = f^{(0)}(1)g^{(0)}(2) \tag{3.146}$$

where f and g are the products of radial factors by spherical harmonics, the former being the variation functions. Zero superscripts indicate that the one-electron wave functions are initial guesses, a concept that will soon be understood, and 1 and 2 inside parentheses stand for spherical coordinates (r_1, θ_1, ϕ_1) and (r_2, θ_2, ϕ_2) of electrons 1 and 2. The average repulsion potential "felt" by electron 1 is obtained by integrating the density of electron 2 over all its positions (integral over $dv_2 = r_2^2 \sin\theta_2 dr_2 d\theta_2 d\phi_2$). In atomic units (see Appendix), this average potential is given by

$$V_1^{(0)}(1) = \int \frac{\left|g^{(0)}(2)\right|^2}{r_{12}} \mathrm{d}v_2 \tag{3.147}$$

(**E31**). The above potential is a function of the coordinates of electron 1 only, since the integration over $\mathrm{d}v_2$ removes the coordinates of electron 2, thus replacing

instantaneous repulsions between electrons 1 and 2 by a fictitious repulsion between the charge of electron 1 and the charge of electron 2 averaged over the entire volume of the coordinates of electron 2. Further integration of (3.147) with respect to the angular coordinates of electron 1, (θ_1, ϕ_1),

$$V_1^{(0)}(r_1) = \frac{\int_0^{2\pi}\int_0^{\pi} V_1^{(0)}(1)\sin\theta_1 \mathrm{d}\theta_1 \mathrm{d}\phi_1}{\int_0^{2\pi}\int_0^{\pi}\sin\theta_1 \mathrm{d}\theta_1 \mathrm{d}\phi_1} \tag{3.148}$$

leads to a function of r_1. Likewise, for electron 2, we can write

$$V_2^{(0)}(2) = \int \frac{\left|f^{(0)}(1)\right|^2}{r_{12}} \mathrm{d}v_1 \tag{3.149}$$

and

$$V_2^{(0)}(r_2) = \frac{\int_0^{2\pi}\int_0^{\pi} V_2^{(0)}(2)\sin\theta_2 \mathrm{d}\theta_2 \mathrm{d}\phi_2}{\int_0^{2\pi}\int_0^{\pi}\sin\theta_2 \mathrm{d}\theta_2 \mathrm{d}\phi_2} \tag{3.150}$$

Having written the one-electron potentials, we can now proceed to write one-electron Schrödinger equations that lead to improved approximations over $f^{(0)}(1)$ and $g^{(0)}(2)$. In atomic units,

$$\left[-\frac{1}{2}\nabla_1^2 - \frac{Z}{r_1} + V_1^{(0)}(r_1)\right] f^{(1)}(1) = \varepsilon_1^{(1)} f^{(1)}(1) \tag{3.151}$$

$$\left[-\frac{1}{2}\nabla_2^2 - \frac{Z}{r_2} + V_2^{(0)}(r_2)\right] g^{(1)}(2) = \varepsilon_2^{(1)} f^{(1)}(2) \tag{3.152}$$

where the second terms inside square brackets are the electron–nucleus attractions (Z is the atomic number), and r_1 and r_2 are the distances from electrons 1 and 2 to the atomic nucleus.

Note that the Hamiltonians for the above equations depend on one-electron wave functions. In particular, $V_1^{(0)}(r_1)$ depends on $g^{(0)}(2)$, and $V_2^{(0)}(r_2)$ depends on $f^{(0)}(1)$. Orbitals $f^{(1)}(1)$ and $g^{(1)}(2)$, obtained by solving these equations, are used to evaluate $V_1^{(1)}$ and $V_2^{(1)}$, and these potential operators are introduced back into the one-electron Schrödinger equations to obtain $f^{(2)}(1)$ and $g^{(2)}(2)$. This iterative procedure continues until no significant change is found in the one-electron wave functions and corresponding energies, from one iteration to the next. Then, one says that self-consistency has been attained. Energy eigenvalues ε_1 and ε_2 include the 1–2 and 2–1 repulsions, respectively. These repulsions are the same, so the sum $\varepsilon_1 + \varepsilon_2$ contains the electron–electron repulsion twice. To obtain the total energy, the

repulsion between electron densities $|f(1)|^2$ and $|g(2)|^2$ has to be subtracted from the sum $\varepsilon_1 + \varepsilon_2$,

$$E = \varepsilon_1 + \varepsilon_2 - \iint \frac{|f(1)|^2 |g(2)|^2}{r_{12}} \mathrm{d}v_1 \mathrm{d}v_2 = \varepsilon_1 + \varepsilon_2 - J_{12} \tag{3.153}$$

where J_{12} is called the **Coulomb integral**.

The above **self-consistent field** (SCF) procedure can be easily generalized for a product of spatial orbitals. However, a simple product wave function is not antisymmetric with respect to the exchange of spatial and spin coordinates of all electrons. Considering an antisymmetrized product of spin-orbitals (a Slater determinant), the total energy includes exchange integrals in addition to Coulomb integrals. The Hartree–Fock total energy for a closed shell configuration of a molecule with n electrons, i.e., $n/2$ doubly occupied molecular orbitals, is given by

$$E_{\mathrm{HF}} = 2 \sum_{i=1}^{n/2} \varepsilon_i - \sum_{i=1}^{n/2} \sum_{j=1}^{n/2} (2J_{ij} - K_{ij}) + V_{\mathrm{NN}} \tag{3.154}$$

(**E32**, **E33**). In atomic units, **Coulomb integrals** are given by

$$J_{ij} = \iint \frac{\phi_i^*(1)\phi_i(1)\phi_j^*(2)\phi_j(2)}{r_{12}} \mathrm{d}v_1 \mathrm{d}v_2 \tag{3.155}$$

and **exchange integrals** by

$$K_{ij} = \iint \frac{\phi_i^*(1)\phi_j(1)\phi_j^*(2)\phi_i(2)}{r_{12}} \mathrm{d}v_1 \mathrm{d}v_2 \tag{3.156}$$

(**E34**, **E35**). While each Coulomb integral represents the repulsion between two orbital densities, each exchange integral can be physically associated with the interaction between two overlap densities and results from antisymmetry in the wave function. Integrations over electron coordinates in (3.155) and (3.156) have the effect of dealing with Coulombic and exchange electron–electron interactions, not as real instantaneous repulsions, but as averages over the spatial coordinates of the electrons. This is the price the Hartree–Fock method pays for dealing with one-electron functions (orbitals) and, at the same time, having to calculate integrals involving the electron–electron distance r_{ij}.

3.9.1 Slater-Type Orbitals

Hartree–Fock calculations on atoms frequently use Slater-type orbitals. In atomic units, a **Slater-type orbital** (STO) is given by

$$N_{n\zeta} r^{n-1} \mathrm{e}^{-\zeta r} Y_{\ell}^{m}(\theta, \phi) \tag{3.157}$$

where n, ℓ, and m are quantum numbers, ζ is the **orbital exponent**, and the normalization constant is given by

$$N_{n\zeta} = \frac{(2\zeta)^{n+1/2}}{[(2n)!]^{1/2}} \tag{3.158}$$

The *Mathematica* code **M7** computes the normalization constants of radial factors of Slater-type orbitals, for $n \leq 3$ (**E36**, **E37**).

3.9.2 Hartree–Fock Equations

When the Hartree–Fock SCF procedure is carried out for a closed shell electronic configuration, that is, for a nondegenerate ground state (electron-spin multiplicity $2S + 1 = 1$), the wave function is given by a single Slater determinant of spin-orbitals. The orthonormal orbitals that minimize the Hartree–Fock energy (3.154) satisfy the **Hartree–Fock equations**,

$$\hat{F}(1)\phi_i(1) = \varepsilon_i \phi_i(1) \tag{3.159}$$

where ε_i is an eigenvalue (an orbital energy), ϕ_i is the corresponding eigenfunction (a spatial orbital), and the **Fock operator** in atomic units is given by

$$\hat{F}(1) = -\frac{1}{2}\nabla_1^2 - \sum_{\alpha} \frac{Z_\alpha}{r_{1\alpha}} + \sum_{j=1}^{n/2} \left[2\hat{J}_j(1) - \hat{K}_j(1)\right] \tag{3.160}$$

where the subscript α runs over the nuclei, and the **Coulomb** and **exchange operators** are defined by the following equalities:

$$\hat{J}_j(1)\phi_i(1) = \phi_i(1)\left[\int \frac{\phi_j^*(2)\phi_j(2)}{r_{12}} \mathrm{d}v_2\right] \tag{3.161}$$

and

$$\hat{K}_j(1)\phi_i(1) = \phi_j(1)\left[\int \frac{\phi_j^*(2)\phi_i(2)}{r_{12}} \mathrm{d}v_2\right] \tag{3.162}$$

Equations (3.159) have to be solved iteratively until self-consistency is attained, since the Fock operator depends on ϕ_j for $j \neq i$ [see (3.160)].

3.9.3 Hartree–Fock–Roothaan Equations

In 1951, Roothaan expanded ϕ_i in one-electron basis functions χ_t centered on the nuclei of the molecule,

$$\phi_i = \sum_{t=1}^{b} \chi_t c_{ti} \tag{3.163}$$

where b is the total number of basis functions. Substitution of (3.163) in the Fock equations (3.159) and multiplication by the complex conjugate of χ_s followed by integration leads to

$$\sum_{t=1}^{b} (F_{st} - \varepsilon_i S_{st}) c_{ti} = 0 \quad s = 1, 2, \ldots, b \tag{3.164}$$

where

$$F_{st} = \int \chi_s^* \hat{F} \chi_t \mathrm{d}v \quad S_{st} = \int \chi_s^* \chi_t \mathrm{d}v \tag{3.165}$$

Equations (3.164) are called **Hartree–Fock–Roothaan equations** and form a system of linear homogeneous equations in which each unknown ε_i corresponds to a set of b coefficients c_{ti} (t = 1, 2, …, b) that define a normalized ϕ_i. The nontrivial solution of (3.164) implies

$$\det(F_{st} - \varepsilon_i S_{st}) = 0 \tag{3.166}$$

This equation leads to a polynomial whose unknowns are the orbital energies ε_i. Solving the equations that result from the substitution of each determined ε_i in (3.164) and normalizing ϕ_i gives the coefficients c_{ti}, where t = 1, 2, …, b. The F matrix elements are obtained from the Fock operator, which in turn depends on the ϕ_i orbitals. Therefore, the Hartree–Fock–Roothaan equations are solved by an iterative process that ends when self-consistency is reached.

Rearranging (3.164) leads to

$$\sum_{t=1}^{b} F_{st} c_{ti} = \sum_{t=1}^{b} S_{st} c_{ti} \varepsilon_i \quad s = 1, 2, \ldots, b \tag{3.167}$$

which suggests the use of matrix formalism. In fact, the first member shows that each row of the F matrix (a fixed s value, with t varying from 1 to b) multiplies a column of the matrix of coefficients (with a defined i, with t varying from 1 to b). Repeating this procedure for all values of s and subsequently all values of i gives the product of the matrices F and C. In the second member of (3.167), the matrix

S replaces the matrix F of the first member and each column of the matrix of coefficients C is multiplied by the corresponding orbital energy value. In matrix formalism, Eq. (3.167) are expressed as

$$\boldsymbol{FC} = \boldsymbol{SC\varepsilon} \tag{3.168}$$

3.9.4 Correlation Energy

A Hartree–Fock SCF wave function takes the interactions between electrons only in an average way, whereas interelectronic repulsions are in reality instantaneous. Thus, the difference between the exact nonrelativistic energy and the Hartree–Fock energy is defined as the **correlation energy**

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}} \tag{3.169}$$

According to the variational theorem (3.126), the exact nonrelativistic energy is lower than the Hartree–Fock energy. Therefore, the correlation energy is always negative.

One apparently obvious method of dealing with electron correlation involves the explicit inclusion of electron–electron distance in the wave function. However, due to the occurrence of many difficult integrals involving the electron–electron distance, this method can be applied only to very small systems, one "historical" example being the helium atom (Coulson and Neilson 1961). Another method of dealing with correlation is **configuration interaction**, whereby the wave function is taken as a linear combination of ground and excited configurations,

$$\Psi = \sum_i c_i \Phi_i \tag{3.170}$$

and each Φ_i corresponds to a single Slater determinant or to a linear combination of a few Slater determinants in open shell configurations. Significant contributions in (3.170) require excited configurations with the same symmetry as the ground state configuration.

For a specified method of calculation and set of basis functions, the correlation correction can be obtained as a percentage using the formula $(E_{\text{calc}} - E_{\text{HF}})/E_{\text{calc}} \times 100$. For small molecules, energy values can be readily found in CCCBDB (see Credits). Considering the B3LYP method and the aug-cc-pVTZ basis set (the student can find the technical details of this method of calculation and basis set explained in Sect. 3.15), the correlation correction is about 0.5 % for H_2O and approximately 0.8 % for CH_4 and H_3CCH_3. In chemical terms, these errors correspond to important energies, roughly 3 to 4 times the energy of one CH bond in CH_4 (435 kJ mol^{-1}).

3.10 Density Functional Theory

3.10.1 Electron Probability Density

Consider the nondegenerate ground state wave function of one molecule, $\Psi_0(s_1, \ldots, s_n)$, where s_k represents the set of three spatial coordinates (x_k, y_k, z_k) and one spin coordinate of electron k ($k = 1, \ldots, n$). This wave function contains $3n$ spatial coordinates of the n electrons. The probability of finding electron 1 in volume dv_1 irrespective of its spin and positions and spins of other electrons is given by

$$dp(\boldsymbol{r}_1) = \left[\int \ldots \int |\Psi_0(s_1, \ldots, s_n)|^2 dv_2 \ldots dv_n d\sigma_1 \ldots d\sigma_n\right] dv_1 \tag{3.171}$$

where the integrals extend to the spatial coordinates of electrons 2 to n and to the spin coordinates of electrons 1 to n. The electrons are indistinguishable, so the probability for finding *any* electron, not just electron 1, in the volume dv, is n times (3.171). Therefore, the ground state **electron probability density** is

$$\rho_0(\boldsymbol{r}) = n\left[\int \ldots \int |\Psi_0(s_1, \ldots, s_n)|^2 dv_2 \ldots dv_n d\sigma_1 \ldots d\sigma_n\right] \tag{3.172}$$

$\rho_0(\boldsymbol{r})$ is a one-electron function central to the density functional theory, as we are going to find out. Its units are volume^{-1} (in atomic units, 1/bohr^3; see Appendix). The **electron charge density** is obtained by multiplying ρ_0 by the electron charge $-e$. Since the electron charge in atomic units is -1, the electron charge density is given in atomic units by $-\rho_0$.

Note that the integral of $\rho_0(\boldsymbol{r})$ over dv [see (3.172); integration of the right-hand side of (3.172) is over dv_1] gives n, the total number of electrons of the molecule, since the ground state wave function is normalized,

$$\int \rho_0(\boldsymbol{r}) dv = n \tag{3.173}$$

3.10.2 External Potential

We now consider the pure electronic Hamiltonian (3.141) and the electronic Schrödinger equation (3.140), within the Born–Oppenheimer approximation (Sect. 3.8). The electronic Schrödinger equation is solved for fixed positions of the nuclei, since nucleus–nucleus repulsions simply add a constant to the electronic Hamiltonian in the electronic Schrödinger equation (3.140). The set of nuclei in their fixed positions in the molecule, that is, the nuclear framework of the molecule, originates on each electron i the potential given by

$$v(\boldsymbol{r}_i) = -\sum_{\alpha} \frac{Z_\alpha}{r_{\alpha i}} \tag{3.174}$$

in atomic units. In density functional theory, (3.174) is called **external potential**, because it reflects the nuclear framework of the molecule, which is external to the system of electrons. Considering that we know the number of electrons of the molecule, it follows that the nuclear framework of the molecule, and consequently, the external potential, determine the molecular Hamiltonian, since the kinetic energy operator and the electron–electron repulsion operator are the molecule's independent operators. The sum of $v(\boldsymbol{r}_i)$ over all electrons gives

$$\hat{V}_{\mathrm{Ne}} = \sum_{i=1}^{n} v(\boldsymbol{r}_i) \tag{3.175}$$

Since the electrons are indistinguishable and $v(\boldsymbol{r})$ commutes with Ψ_0^* [$v(\boldsymbol{r})$ is simultaneously an operator and a function, and as such, commutes with another function], we can write

$$\begin{aligned} &\int \ldots \int \Psi_0^* \left(\sum_{i=1}^{n} v(\boldsymbol{r}_i) \right) \Psi_0 dv_1 \ldots \mathrm{d}v_n \mathrm{d}\sigma_1 \ldots \mathrm{d}\sigma_n \\ &= \int v(\boldsymbol{r}) n \left[\int \ldots \int |\Psi_0|^2 \mathrm{d}v_2 \ldots \mathrm{d}v_n \mathrm{d}\sigma_1 \ldots \mathrm{d}\sigma_n \right] \mathrm{d}v \end{aligned} \tag{3.176}$$

Using (3.172), we obtain the average value of the nuclei–electron attractions

$$V_{\mathrm{Ne}}[\rho_0] = \int v(\boldsymbol{r}) \rho_0(\boldsymbol{r}) \mathrm{d}v \tag{3.177}$$

3.10.3 Functional Derivative

We now derive an expression that enables us to evaluate functional derivatives, and in particular, obtain the functional derivative of V_{Ne} with respect to ρ. We begin by considering the set of all continuous plane curves connecting two given points a and b, and the length associated with each curve. This is a simple example of a functional, a rule that associates a number, in this case the length of each curve or functional value, with a function, in this example each curve connecting points a and b. The length of each plane curve is given by the following integral:

$$J = \int_a^b \mathrm{d}s = \int_a^b \left(\mathrm{d}x^2 + \mathrm{d}y^2 \right)^{1/2} = \int_a^b \left(1 + y'^2 \right)^{1/2} \mathrm{d}x \tag{3.178}$$

where $dy = y'dx$. This integral sets a correspondence between its value J, a real number, and the derivative of each plane curve connecting the given points a and b, the function y'. The integral J is a functional of y', and this is represented by $J[y']$. We are interested in finding the curve with the shortest length. Of course, this is a trivial question, since the shortest path between two points is a straight line. However, solving this problem in a systematic way requires evaluation of functional derivatives, an important question in density functional theory. To this end, let us consider the more general functional

$$J[y] = \int_a^b F(x, y, y')\mathrm{d}x \tag{3.179}$$

and an increment $h(x)$ to the function $y(x)$,

$$y(x) \Rightarrow y(x) + h(x) \tag{3.180}$$

so that $h(a) = h(b) = 0$ (a and b are fixed endpoints, so the variations at these points are zero). The corresponding increment for the functional (3.179) is

$$\Delta J = J[y+h] - J[y] = \int_a^b [F(x, y+h, y'+h') - F(x, y, y')]\mathrm{d}x \tag{3.181}$$

and the Taylor expansion of the integrand gives

$$\Delta J = \int_a^b \left[F_y h + F_{y'} h'\right]\mathrm{d}x + \cdots \tag{3.182}$$

where F_y and $F_{y'}$ represent partial derivatives of F with respect to y and y', and the dots represent terms of order higher than 1, relative to h and h'. The differential of $J[y]$ is

$$\delta J = \int_a^b \left[F_y h + F_{y'} h'\right]\mathrm{d}x \tag{3.183}$$

(note the use of the symbol δ to represent a functional differential and compare (3.183) with (3.179)). Integration by parts of the second term inside square brackets leads to

$$\begin{aligned}\delta J &= \int_a^b \left[F_y h + F_{y'} h'\right]\mathrm{d}x = \int_a^b F_y h\mathrm{d}x + F_{y'} h\Big|_a^b - \int_a^b \left(\frac{\mathrm{d}}{\mathrm{d}x} F_{y'}\right) h\mathrm{d}x \\ &= \int_a^b \left(F_y - \frac{\mathrm{d}}{\mathrm{d}x} F_{y'}\right) h\mathrm{d}x\end{aligned} \tag{3.184}$$

since the second term is zero because $h(a) = h(b) = 0$. At a stationary point,

$$\delta J = \int_a^b \left(F_y - \frac{\mathrm{d}}{\mathrm{d}x} F_{y'} \right) h(x) \mathrm{d}x = 0 \tag{3.185}$$

Since $h(x)$ is arbitrary, we can conclude that

$$F_y - \frac{\mathrm{d}}{\mathrm{d}x} F_{y'} = 0 \tag{3.186}$$

This equation, called the **Euler–Lagrange differential equation**, is central to the calculus of variations. The necessary condition for $J[y]$ to have an extremum (maximum or minimum) is that $y(x)$ satisfy the Euler–Lagrange equation (see *Gelfand and Fomin*, Further Reading, Chap. 1, Sect. 4).

Returning now to the example of the curve in a plane with the shortest length between two points a and b and applying the Euler–Lagrange equation, we obtain $F_y = 0$, $F_{y'} = y'/\sqrt{(1 + y'^2)}$, $\mathrm{d}F_{y'}/\mathrm{d}x = 0$, and $y' =$ constant, that is, $y(x) = \mathrm{A}x + \mathrm{B}$, as anticipated.

Since $h(x)$ is an arbitrarily small variation in $y(x)$ [see (3.180)], we can represent $h(x)$ by δy and rewrite the last equality of (3.184) as

$$\delta J = \int_a^b \left(F_y - \frac{\mathrm{d}}{\mathrm{d}x} F_{y'} \right) \delta y \mathrm{d}x \tag{3.187}$$

The **functional or variational derivative** of J is given by

$$\frac{\delta J}{\delta y} = F_y - \frac{\mathrm{d}}{\mathrm{d}x} F_{y'} \tag{3.188}$$

(a proof of this statement can be found in *Gelfand and Fomin*, Further Reading, Chap. 1, Sect. 7). Hence, the Euler–Lagrange equation (3.186) can be formulated as expressing that the functional derivative at a stationary point is zero, a result that is also applicable to an ordinary derivative of a function at a stationary point.

We now consider (3.177) for a trial electron density ρ and, using (3.179), write $F(\boldsymbol{r}, \rho) = v(\boldsymbol{r})\rho(\boldsymbol{r})$ and apply (3.188). Since F does not explicitly depend on ρ', its derivative with respect to ρ' is zero. In turn, the derivative of F with respect to ρ is $v(\boldsymbol{r})$. Hence, using (3.188), we conclude that the functional derivative of V_{Ne} with respect to ρ is equal to the external potential $v(\boldsymbol{r})$,

$$\frac{\delta V_{\mathrm{Ne}}[\rho(\boldsymbol{r})]}{\delta \rho(\boldsymbol{r})} = v(\boldsymbol{r}) \tag{3.189}$$

This result connects the external potential $v(\boldsymbol{r})$ with the average nuclei–electrons attraction potential V_{Ne}. The *Mathematica* code **M8** calculates several functional derivatives and confirms (3.189).

3.10.4 Hohenberg–Kohn Theorems

The ground state wave function of a molecule is a function of $4n$ electron coordinates ($3n$ spatial coordinates and n spin coordinates), whereas the electron probability density is a function of only three spatial coordinates. Unlike the wave function, the electron probability density of a molecule is an experimentally determined observable that can be obtained from X-ray diffraction intensities of molecular crystals.

For a nondegenerate ground state, Pierre Hohenberg and Walter Kohn demonstrated in 1964 that the ground state electron probability density ρ_0 *uniquely* determines the Hamiltonian operator, the ground state molecular energy, and all other ground state molecular properties (Hohenberg and Kohn 1964). This theorem was proved by *reductio ad absurdum*, that is, by showing that its contradiction would lead to a false conclusion. In particular, it was assumed that the same ground state electron density ρ_0 can give rise to two external potentials differing by more than a constant, and consequently, to two different n-electron Hamiltonians. This premise leads to a logical impossibility, thus proving that the ground state electron probability density ρ_0 *uniquely* determines the Hamiltonian operator, the ground state molecular energy, and all other ground state molecular properties. The ground state electron density ρ_0 determines $v(\boldsymbol{r})$, which is a system-dependent operator, since it involves the nuclear framework of the molecule. In turn, the external potential determines the electronic Hamiltonian, because the kinetic energy operator and the electron–electron repulsion operator are universal operators, dependent only on the number of electrons of the molecule.

The electronic Hamiltonian is the sum of the operators for the kinetic energy of electrons and the potential energy for the electron–nucleus attractions and electron–electron repulsions [see (3.141)]. Consequently, we can write

$$E_0 = T_{\mathrm{e}} + V_{\mathrm{Ne}} + V_{\mathrm{ee}} \tag{3.190}$$

where E_0, T_{e}, V_{Ne}, and V_{ee} represent ground state averages of the Hamiltonian, electron kinetic energy operator, electron–nucleus attractions, and electron–electron repulsions, respectively. Since E_0 is a functional of the electron density ρ_0 and V_{Ne} is a known functional of ρ_0 [see (3.177)], we conclude that T_{e} and V_{ee} are also functionals of ρ_0 and write

$$E_0[\rho_0] = T_{\mathrm{e}}[\rho_0] + \int v(\boldsymbol{r})\rho_0(\boldsymbol{r})\mathrm{d}v + V_{\mathrm{ee}}[\rho_0] \tag{3.191}$$

Note that the Hohenberg–Kohn theorem does not tell us how to calculate E_0 from ρ_0, since T_e and V_{ee} are unknown functionals of ρ_0.

For a trial electron density $\rho(\boldsymbol{r})$ that satisfies

$$\int \rho(\boldsymbol{r}) \mathrm{d}v = n \tag{3.192}$$

and $\rho(\boldsymbol{r}) \geq 0$ for all $\boldsymbol{r}$, Hohenberg and Kohn also proved that

$$E[\rho] = T_e[\rho] + \int v(\boldsymbol{r})\rho(\boldsymbol{r})\mathrm{d}v + V_{ee}[\rho] \geq E_0[\rho_0] \tag{3.193}$$

that is, the true ground state electron density minimizes the energy functional. This variational theorem shows that no trial electron density can give a lower ground state energy than the exact ground state electron density. Note that the validity of this theorem requires the use of the electronic Hamiltonian, as mentioned above [see (3.190)].

3.10.5 Kohn–Sham Method

Let us now recall the first Hohenberg–Kohn theorem, which tells us that the ground state electron probability density ρ_0 *uniquely* determines the ground state molecular energy and all other ground state molecular properties. However, this theorem does not tell us how to calculate E_0 from ρ_0, since T_e and V_{ee} in (3.191) are unknown functionals of ρ_0, nor does it provide us with a method for evaluating ρ_0 without having first to obtain the ground state wave function.

In 1965, Kohn and Sham considered a fictitious reference system (we label this system by the subscript s) formed by n noninteracting electrons, with each electron being subject to the effective potential $v_s(\boldsymbol{r}_i)$, which was chosen so that the ground state electron density of the reference system ρ_s is equal to the exact ground state electron density of the real molecule, $\rho_s = \rho_0$ (Kohn and Sham 1965). Note that the electron probability density is a function of 3 variables and is obtained from a wave function of 4n variables by integration over 4n-3 variables [see (3.172)]. The connection between this noninteracting system and the target molecule is provided by the first Hohenberg–Kohn theorem, since the ground state electron density of the reference system ρ_s, being equal to ρ_0, uniquely determines the ground state molecular energy and all other ground state molecular properties of our real target system of interacting electrons. The noninteracting electrons of the fictitious system enable us to define one-electron wave functions (spin-orbitals) and form a Slater determinant of the lowest-energy Kohn–Sham spin-orbitals,

$$\psi_{s,0} = \left|\chi_1^{KS}\chi_2^{KS}\ldots\chi_n^{KS}\right| \quad \chi_k^{KS} = \theta_i^{KS}\alpha \quad \chi_{k+1}^{KS} = \theta_i^{KS}\beta \tag{3.194}$$

where k varies from 1 to n, i varies from 1 to $n/2$ (for n electrons, the closed shell configuration has $n/2$ occupied orbitals), and each spin-orbital χ_k is the product of a **Kohn–Sham orbital** θ_i and the spin function α or β. Note that the Slater

determinant includes the normalizing factor $1/\sqrt{n!}$. Just as the orthonormal orbitals that minimize the Hartree–Fock energy satisfy the Fock equations, the Kohn–Sham orthonormal orbitals that minimize the ground state energy satisfy the Kohn–Sham eigenvalue equations

$$\hat{h}^{\mathrm{KS}}(1)\theta_i^{\mathrm{KS}}(1) = \varepsilon_i^{\mathrm{KS}}\theta_i^{\mathrm{KS}}(1) \tag{3.195}$$

where the one-electron Kohn–Sham Hamiltonian in atomic units is given by

$$\hat{h}^{\mathrm{KS}}(1) = \hat{T}_{\mathrm{e}}(\boldsymbol{r}_1) + v_s(\boldsymbol{r}_1) \tag{3.196}$$

and the electron density of the Slater determinant of spin-orbitals is

$$\rho_0 = \rho_s = \sum_{i=1}^{n/2} \left|\theta_i^{\mathrm{KS}}\right|^2 \tag{3.197}$$

(**E38**).

We now recall Eq. (3.191), which represents the first Hohenberg–Kohn theorem,

$$E_0[\rho_0] = T_{\mathrm{e}}[\rho_0] + \int v(\boldsymbol{r})\rho_0(\boldsymbol{r})\mathrm{d}v + V_{\mathrm{ee}}[\rho_0] \tag{3.198}$$

with T_{e} and V_{ee} unknown functionals of ρ_0. The kinetic energy operator is a sum of one-electron kinetic energy operators. Hence, for the noninteracting system of Kohn and Sham, it can be shown that the kinetic energy is given by

$$\begin{aligned} T_{e,s}[\rho_0] &= \int \ldots \int \psi_{s,0}^* \left[\sum_{i=1}^{n} \hat{T}_{\mathrm{e}}(i)\right] \psi_{s,0}\mathrm{d}v_1 \ldots \mathrm{d}v_n \mathrm{d}\sigma_1 \ldots \mathrm{d}\sigma_n \\ &= \sum_{i=1}^{n} \int (\theta_i^{\mathrm{KS}})^*(1)\hat{T}_{\mathrm{e}}(1)\,\theta_i^{\mathrm{KS}}(1)\mathrm{d}v_1 \end{aligned} \tag{3.199}$$

(**E39**). Note that this kinetic energy is not equal to the kinetic energy of the real interacting system, even if the noninteracting and interacting systems share the same electron density. Therefore, we can write

$$T_{\mathrm{e}}[\rho_0] = T_{e,s}[\rho_0] + \Delta T_{\mathrm{e}}[\rho_0] \tag{3.200}$$

where ΔT_{e} is the difference between the average kinetic energy of the molecule and the average kinetic energy of the reference system. We now consider the third term on the right-hand side of (3.198) and write

$$V_{\mathrm{ee}}[\rho_0] = \frac{1}{2}\iint \frac{\rho_0(\boldsymbol{r}_1)\rho_0(\boldsymbol{r}_2)}{r_{12}}\mathrm{d}v_1\mathrm{d}v_2 + \Delta V_{\mathrm{ee}}[\rho_0] \tag{3.201}$$

where the second term of the right-hand side of this equation is the difference between the average electron–electron repulsion energy of the molecule and the electrostatic electron–electron repulsion obtained from the ground state electron density ρ_0. In order to understand the 1/2 factor of the first term of (3.201), we need to go back to (3.172) and find that each ρ_0 factor involves a multiplication by n, that is, it runs over n electron labels that are different in both ρ_0 factors, as required by the operator $1/r_{ij}$ ($i \neq j$). For example, for electrons with labels 1, 2, and 3, the six-dimensional integration of (3.201) applies to a sum over electron pairs 1–2, 1–3, 2–1, 2–3, 3–1, and 3–2. Hence, each electron–electron pair is repeated, and the factor 1/2 corrects for that double counting. Considering now the physical meaning of the electron charge density in atomic units $-\rho_0$, the products $-\rho_0(\boldsymbol{r}_1)\mathrm{d}v_1$ and $-\rho_0(\boldsymbol{r}_2)\mathrm{d}v_2$ in the six-dimensional integral of (3.201) represent electronic charges in the volume elements $\mathrm{d}v_1$ and $\mathrm{d}v_2$. Therefore, on the right hand-side of (3.201), the first term represents the electrostatic electron–electron repulsion energy.

Substitution of the result of (3.199) in (3.200) and subsequent substitution of the resulting Eq. (3.200) and of (3.201) in (3.198) leads to

$$\begin{aligned} E_0[\rho_0] = &\sum_{i=1}^{n} \int (\theta_i^{\mathrm{KS}})^*(1)\hat{T}_{\mathrm{e}}(1)\theta_i^{\mathrm{KS}}(1)\mathrm{d}v_1 \\ &+ \int v(r)\rho_0(r)\mathrm{d}v + \frac{1}{2}\iint \frac{\rho_0(\boldsymbol{r}_1)\rho_0(\boldsymbol{r}_2)}{r_{12}}\mathrm{d}v_1\mathrm{d}v_2 + E_{xc}[\rho_0] \end{aligned} \tag{3.202}$$

where the delta terms ΔT_{e} and ΔV_{ee} are grouped in the so-called **exchange-correlation energy functional** defined by

$$E_{xc}[\rho_0] \equiv \Delta T_{\mathrm{e}}[\rho_0] + \Delta V_{\mathrm{ee}}[\rho_0] \tag{3.203}$$

The energy components of $E_{xc}[\rho_0]$ are the **kinetic correlation energy**, which is the ΔT_{e} difference between the kinetic energy for the real molecule and the kinetic energy of the reference system, the **exchange energy**, which results from the antisymmetric characteristic of the wave function [see (3.194)], the **Coulombic correlation energy**, which accounts for the influence of electron correlation on the electrostatic electron–electron repulsions described by the first term of the right-hand side of (3.201), and the **self-interaction correction**, which results from the fact that in the electrostatic electron–electron repulsion energy, the electron density $\rho_0(\boldsymbol{r}_1)$ is different from zero in $\mathrm{d}v_1$, and so electron 1 erroneously interacts with itself, the same thing happening with $\rho_0(\boldsymbol{r}_2)$ and $\mathrm{d}v_2$.

Since the exact form of the exchange-correlation energy functional is unknown, one needs to resort to approximate functionals to describe it. Having arrived at the expression for the ground state energy functional [see (3.202)], we can now be a bit more explicit about the external potential operator of (3.196) for the one-electron Kohn–Sham Hamiltonian

$$\hat{h}^{\mathrm{KS}}(1) = \hat{T}_{\mathrm{e}}(\boldsymbol{r}_1) + v_s(\boldsymbol{r}_1) = -\frac{1}{2}\nabla_1^2 - \sum_{\alpha}\frac{Z_\alpha}{r_{1\alpha}} + \int \frac{\rho(\boldsymbol{r}_2)}{r_{12}}\mathrm{d}v_2 + v_{xc}(1) \quad (3.204)$$

where $v_{xc}(\boldsymbol{r})$ is the **one-electron exchange-correlation potential**, defined by the derivative of the exchange-correlation energy functional,

$$v_{xc}(\boldsymbol{r}) \equiv \frac{\delta E_{xc}[\rho_0(\boldsymbol{r})]}{\delta \rho_0(\boldsymbol{r})} \quad (3.205)$$

In analogy with (3.189), this definition implies that

$$E_{xc}[\rho_0] = \int v_{xc}(\boldsymbol{r})\rho_0(\boldsymbol{r})dv \quad (3.206)$$

[see (3.177)]. This equation results from (3.205), since we defined $v_{xc}(\boldsymbol{r})$ from $E_{xc}[\rho_0]$, not the other way round.

3.10.6 Overview

Let us now summarize the main steps in the derivation of the density functional theory equations. The most important statement lies in the first Hohenberg–Kohn theorem, which says that the ground state electron probability density ρ_0 *uniquely* determines the Hamiltonian operator, the ground state molecular energy, and all other ground state molecular properties.

Within the Born–Oppenheimer approximation, the electronic Hamiltonian is the sum of operators for the kinetic energy of electrons, potential energy of the electron–nucleus attractions, and potential energy of electron–electron repulsions. Based on the first Hohenberg–Kohn theorem, we conclude that the electronic energy and its component terms are functionals of ρ_0. However, this theorem does not tell us how to calculate E_0 from ρ_0, since T_{e} and V_{ee} are unknown functionals of ρ_0.

The Kohn–Sham method overcomes this difficulty by considering a fictitious reference system formed by n noninteracting electrons, with each electron subject to an unknown effective potential, which in turn is supposed to originate the exact ground state electron density for the real molecule.

The fictitious reference system of noninteracting electrons leads to one-electron functions (i.e., spin-orbitals), thus enabling the calculation of the average value of the kinetic operator for the corresponding Slater determinant. Since this kinetic energy is not equal to the kinetic energy of the real interacting system, one needs to determine the difference ΔT_{e} between the average kinetic energy of the real molecule and the average kinetic energy of the reference system.

In turn, the average electron–electron repulsion energy is another unknown functional. However, since the electrostatic electron–electron repulsion obtained from the ground state electron density ρ_0 can in principle be calculated, the

difference ΔV_{ee} between the average electron–electron repulsion energy of the real molecule and the electrostatic electron–electron repulsion of the fictitious noninteracting system remains to be calculated.

Now the delta terms ΔT_e and ΔV_{ee} are grouped in the so-called exchange-correlation energy functional, whose exact form is unknown. For this reason, in the various DFT methods of calculation, approximate functionals are used to describe the exchange-correlation energy functional.

One practical question remains to be answered: how to obtain the electron density of the target molecule? The answer to this question is provided by the variational Hohenberg–Kohn theorem, which shows that the true ground state electron density minimizes the energy functional obtained using the exact electronic Hamiltonian. This theorem suggests a way of improving the trial electron density, obtained as the sum of the squares of the absolute values of the spatial Kohn–Sham orbitals [see (3.197)], by varying the electron density, so as to minimize the energy functional. Like the one-electron Hartree–Fock equations (3.159), the one-electron Kohn–Sham equations (3.195) also have to be solved iteratively, since the one-electron Kohn–Sham Hamiltonian depends on all one-electron θ_i functions [see (3.197) and (3.204)].

Note the parallelism between the Hartree–Fock and Kohn–Sham methods. Both assume fictitious systems in order to define one-electron functions (orbitals). However, in the Hartree–Fock method, each electron is subject to the average Coulomb repulsion of the remaining electrons, to which is added the exchange potential that results from the antisymmetry requirement. In turn, in the Kohn–Sham method, a fictitious reference system is formed by n noninteracting electrons, with each electron being subject to the effective potential $v_s(\boldsymbol{r}_i)$ so chosen that the ground state electron density of the reference system ρ_s is equal to the exact ground state electron density of the real molecule, $\rho_s = \rho_0$. The Kohn–Sham equations (3.195) and (3.196) correspond to the Hartree–Fock equations (3.159) and (3.160). Both Hartree–Fock and Kohn–Sham methods can be expressed by matrix formalisms, whereby Hartree–Fock orbitals are replaced by Kohn–Sham orbitals, and the Fock matrix is replaced by a Kohn–Sham matrix.

3.11 Perturbation Theory

After the variational method, the major method for approximately solving the Schrödinger equation is perturbation theory, applicable to a system B whose Schrödinger equation cannot be solved but can be obtained from the Schrödinger equation for system A by the introduction of a small perturbation. A hydrogen atom in the presence of a small uniform electric field illustrates the use of perturbation theory. In this section, both the unperturbed Hamiltonian (system A) and the perturbation are time-independent operators.

3.11.1 Nondegenerate Energy Level

Consider the time-independent Schrödinger equation

$$\hat{H}\psi_k = E_k\psi_k \tag{3.207}$$

where ψ_k represents the eigenfunctions of the perturbed system (system B) and assume that we can solve the Schrödinger equation for the unperturbed system (system A)

$$\hat{H}^0\psi_k^{(0)} = E_k^{(0)}\psi_k^{(0)} \tag{3.208}$$

where $\psi_k^{(0)}$ represents the wave function for the unperturbed nondegenerate level with energy $E_k^{(0)}$. Suppose that $\hat{H}$ differs slightly from $\hat{H}^0$ by the perturbation operator $\hat{H}'$,

$$\hat{H} = \hat{H}^0 + \hat{H}' \tag{3.209}$$

We introduce the scaling parameter λ so that the perturbation is applied gradually and there is a continuous change from the unperturbed to the perturbed system. Therefore,

$$\hat{H} = \hat{H}^0 + \lambda\hat{H}' \tag{3.210}$$

where λ varies between 0, for the unperturbed system, and 1, for the perturbed system when the perturbation is fully "switched on." We now consider a nondegenerate eigenfunction $\psi_0^{(0)}$ and corresponding eigenvalue $E_0^{(0)}$ of the unperturbed Hamiltonian $\hat{H}^0$,

$$\hat{H}^0\psi_0^{(0)} = E_0^{(0)}\psi_0^{(0)} \tag{3.211}$$

and the Schrödinger equation for the system with the perturbation fully "switched on,"

$$\hat{H}\psi_0 = E_0\psi_0 \tag{3.212}$$

Substitution of (3.210) in (3.212) shows that both the eigenfunction and the eigenvalue depend on λ,

$$\psi_0 = \psi_0(s,\lambda) \quad E_0 = E_0(\lambda) \tag{3.213}$$

where s represents the space and spin coordinates. We can now expand in Taylor series the eigenfunction and the eigenvalue in powers of λ, obtaining

$$\psi_0 = (\psi_0)_{\lambda=0} + \lambda\left(\frac{\partial\psi_0}{\partial\lambda}\right)_{\lambda=0} + \frac{\lambda^2}{2!}\left(\frac{\partial^2\psi_0}{\partial\lambda^2}\right)_{\lambda=0} + \cdots = \psi_0^{(0)} + \lambda\psi_0^{(1)} + \lambda^2\psi_0^{(2)} + \cdots \tag{3.214}$$

$$E_0 = (E_0)_{\lambda=0} + \lambda\left(\frac{\partial E_0}{\partial\lambda}\right)_{\lambda=0} + \frac{\lambda^2}{2!}\left(\frac{\partial^2 E_0}{\partial\lambda^2}\right)_{\lambda=0} + \ldots = E_0^{(0)} + \lambda E_0^{(1)} + \lambda^2 E_0^{(2)} + \cdots \tag{3.215}$$

where

$$\psi_0^{(k)} \equiv \frac{1}{k!}\left(\frac{\partial^k\psi_0}{\partial\lambda^k}\right)_{\lambda=0} \quad E_0^{(k)} \equiv \frac{1}{k!}\left(\frac{\partial^k E_0}{\partial\lambda^k}\right)_{\lambda=0} \tag{3.216}$$

We assume that

$$\left\langle \psi_0^{(0)} \middle| \psi_0 \right\rangle = 1 \tag{3.217}$$

and expand ψ_0 using (3.214) to obtain

$$\left\langle \psi_0^{(0)} \middle| \psi_0 \right\rangle = 1 = \underbrace{\left\langle \psi_0^{(0)} \middle| \psi_0^{(0)} \right\rangle}_{1} + \lambda \underbrace{\left\langle \psi_0^{(0)} \middle| \psi_0^{(1)} \right\rangle}_{0} + \lambda^2 \underbrace{\left\langle \psi_0^{(0)} \middle| \psi_0^{(1)} \right\rangle}_{0} + \cdots \tag{3.218}$$

Substitution of (3.210), (3.214), and (3.215) in (3.212) gives

$$\begin{aligned}(\hat{H}_0 + \lambda\hat{H}')\left(\psi_0^{(0)} + \lambda\psi_0^{(1)} + \lambda^2\psi_0^{(2)} + \cdots\right) \\ = \left(E_0^{(0)} + \lambda E_0^{(1)} + \lambda^2 E_0^{(2)} + \cdots\right)\left(\psi_0^{(0)} + \lambda\psi_0^{(1)} + \lambda^2\psi_0^{(2)} + \cdots\right)\end{aligned} \tag{3.219}$$

We now equate the coefficients of equal powers of λ and obtain

$$\lambda^0: \quad \left(\hat{H}^0 - E_0^{(0)}\right)\psi_0^{(0)} = 0 \tag{3.220}$$

$$\lambda^1: \quad \left(\hat{H}^0 - E_0^{(0)}\right)\psi_0^{(1)} + \left(\hat{H}' - E_0^{(1)}\right)\psi_0^{(0)} = 0 \tag{3.221}$$

$$\lambda^2: \quad \left(\hat{H}^0 - E_0^{(0)}\right)\psi_0^{(2)} + \left(\hat{H}' - E_0^{(1)}\right)\psi_0^{(1)} - E_0^{(2)}\psi_0^{(0)} = 0 \tag{3.222}$$

Equation (3.220), the unperturbed Schrödinger equation, is automatically satisfied [see (3.211)]. We now multiply the other two equations on the left by the complex conjugate of $\psi_0^{(0)}$ and integrate over all space. Since $\hat{H}^0$ is a Hermitian operator, we obtain the first- and second-order corrections to the energy

$$E_0^{(1)} = \left\langle \psi_0^{(0)} \middle| \hat{H}' \middle| \psi_0^{(0)} \right\rangle \quad E_0^{(2)} = \left\langle \psi_0^{(0)} \middle| \hat{H}' \middle| \psi_0^{(1)} \right\rangle \tag{3.223}$$

We now expand the unknown function $\psi_0^{(1)}$ in the complete set of orthonormal eigenfunctions of the unperturbed Hamiltonian,

$$\psi_0^{(1)} = \sum_{k \neq 0} a_k \psi_k^{(0)} = \sum_{k \neq 0} \left\langle \psi_k^{(0)} \middle| \psi_0^{(1)} \right\rangle \psi_k^{(0)} \tag{3.224}$$

since the unperturbed and perturbed systems are assumed to have the same boundary conditions. The expansion coefficient a_0 is zero because $\psi_0^{(1)}$ and $\psi_0^{(0)}$ are orthogonal [see (3.218)]. Multiplying (3.221) on the left by the complex conjugate of $\psi_k^{(0)}$ and integrating over all space gives

$$\begin{gathered} \left(E_k^{(0)} - E_0^{(0)}\right) \underbrace{\left\langle \psi_k^{(0)} \middle| \psi_0^{(1)} \right\rangle}_{a_k} = -\left\langle \psi_k^{(0)} \middle| \hat{H}' \middle| \psi_0^{(0)} \right\rangle \quad \therefore \\ a_k = \left\langle \psi_k^{(0)} \middle| \psi_0^{(1)} \right\rangle = -\frac{\left\langle \psi_k^{(0)} \middle| \hat{H}' \middle| \psi_0^{(0)} \right\rangle}{E_k^{(0)} - E_0^{(0)}} \end{gathered} \tag{3.225}$$

Hence,

$$\psi_0^{(1)} = -\sum_{k \neq 0} \frac{\left\langle \psi_k^{(0)} \middle| \hat{H}' \middle| \psi_0^{(0)} \right\rangle}{E_k^{(0)} - E_0^{(0)}} \psi_k^{(0)} \tag{3.226}$$

Substitution of (3.226) in (3.223) leads to

$$E_0^{(2)} = -\sum_{k \neq 0} \frac{\left\langle \psi_0^{(0)} \middle| \hat{H}' \middle| \psi_k^{(0)} \right\rangle \left\langle \psi_k^{(0)} \hat{H}' \middle| \psi_0^{(0)} \right\rangle}{E_k^{(0)} - E_0^{(0)}} \tag{3.227}$$

thus enabling us to conclude that the second-order energy can be derived from the first-order correction to the wave function. Also, it can be shown that the third-order energy could also be obtained from the first-order wave function, and in general, the $(2n + 1)$th energy correction can be derived from the nth-order perturbation wave function.

3.11.2 Variational Perturbation Method

Perturbation expressions like (3.226) and (3.227) that involve infinite sums are not suitable for quantitative evaluations. However, the variational principle can be used

to provide an upper bound on the second-order perturbation energy. To this end, multiplying (3.221) on the left by the complex conjugate of $\psi_0^{(1)}$ and integrating over all space leads to

$$\left\langle \psi_0^{(1)} \middle| \hat{H}^0 - E_0^{(0)} \middle| \psi_0^{(1)} \right\rangle = -\left\langle \psi_0^{(1)} \middle| \hat{H}' - E_0^{(1)} \middle| \psi_0^{(0)} \right\rangle \tag{3.228}$$

We now multiply (3.222) on the left by the complex conjugate of $\psi_0^{(0)}$, integrate over all space, and apply the orthogonality relationships of (3.218). Using the Hermitian characteristics of the Hamiltonian and the perturbation operator gives

$$-\left\langle \psi_0^{(1)} \middle| \hat{H}' - E_0^{(1)} \middle| \psi_0^{(0)} \right\rangle = -E_0^{(2)} \tag{3.229}$$

Combining (3.228) and (3.229) allows us to write

$$\left\langle \psi_0^{(1)} \middle| \hat{H}^0 - E_0^{(0)} \middle| \psi_0^{(1)} \right\rangle = -E_0^{(2)} \tag{3.230}$$

We consider now a guess function ϕ and express the variational principle Eq. (3.126) for $\phi - \psi_0^{(1)}$ as

$$\left\langle \phi - \psi_0^{(1)} \middle| \hat{H}^0 - E_0^{(0)} \middle| \phi - \psi_0^{(1)} \right\rangle \geq 0 \tag{3.231}$$

This expression gives

$$\begin{aligned} &\left\langle \phi \middle| \hat{H}^0 - E_0^{(0)} \middle| \phi \right\rangle - \left\langle \phi \middle| \hat{H}^0 - E_0^{(0)} \middle| \psi_0^{(1)} \right\rangle \\ &\quad - \left\langle \psi_0^{(1)} \middle| \hat{H}^0 - E_0^{(0)} \middle| \phi \right\rangle + \left\langle \psi_0^{(1)} \middle| \hat{H}^0 - E_0^{(0)} \middle| \psi_0^{(1)} \right\rangle \geq 0 \end{aligned} \tag{3.232}$$

Substitution of (3.230) in (3.232) leads to

$$\left\langle \phi \middle| \hat{H}^0 - E_0^{(0)} \middle| \phi \right\rangle - \left\langle \phi \middle| \hat{H}^0 - E_0^{(0)} \middle| \psi_0^{(1)} \right\rangle - \left\langle \psi_0^{(1)} \middle| \hat{H}^0 - E_0^{(0)} \middle| \phi \right\rangle \geq E_0^{(2)} \tag{3.233}$$

We now return again to Eq. (3.221), multiply it on the left by the complex conjugate of ϕ, integrate, use the Hermitian characteristics of $\hat{H}^0$, and substitute in (3.233), obtaining

$$\left\langle \phi \middle| \hat{H}^0 - E_0^{(0)} \middle| \phi \right\rangle + \left\langle \phi \middle| \hat{H}' - E_0^{(1)} \middle| \psi_0^{(0)} \right\rangle + \left\langle \psi_0^{(0)} \middle| \hat{H}' - E_0^{(1)} \middle| \phi \right\rangle \geq E_0^{(2)} \tag{3.234}$$

The first member of this equation is a functional of ϕ, $J[\phi]$, which is an upper bound to the second-order perturbation energy (**E40**).

3.11.3 Degenerate Energy Level

An energy level with a degeneracy degree d has d linearly independent wave functions,

$$\psi_{0,1}^{(0)}, \psi_{0,2}^{(0)}, \ldots, \psi_{0,d}^{(0)} \tag{3.235}$$

which are eigenfunctions of the unperturbed Hamiltonian,

$$\left(\hat{H}^0 - E_0^{(0)}\right)\psi_{0,k}^{(0)} = 0 \tag{3.236}$$

Note that labels 1, 2, …, d do not necessarily imply that the corresponding states are the lowest-lying states. The wave functions (3.235) are assumed to be orthonormal. Every linear combination of them is still an eigenfunction of the unperturbed Hamiltonian with the same eigenvalue,

$$\left(\hat{H}^0 - E_0^{(0)}\right)\sum_{k=1}^{d} a_k \psi_{0,k}^{(0)} = 0 \tag{3.237}$$

Assuming that the perturbation splits the degeneracy of the energy level, there is a unique normalized eigenfunction ψ_0 with eigenvalue E_0 for the total Hamiltonian,

$$\left(\hat{H} - E_0\right)\psi_0 = 0 \tag{3.238}$$

However, there is a problem in defining $\lim_{\lambda\to 0}\psi_0$, since there is an infinite number of linear combinations of the wave functions (3.235) that satisfy the Schrödinger equation for the unperturbed Hamiltonian and correspond to the *same* energy. Let us assume that

$$\lim_{\lambda\to 0}\psi_0 = \phi_0^{(0)} = \sum_{k=1}^{d} a_k \psi_{0,k}^{(0)} \tag{3.239}$$

How can we determine the coefficients a_k? As previously done for the nondegenerate energy level, the eigenfunction ψ_0 and the eigenvalue E_0 can be expanded as Taylor series in powers of λ, yielding

$$\psi_0 = \phi_0^{(0)} + \lambda\psi_0^{(1)} + \lambda^2\psi_0^{(2)} + \ldots \tag{3.240}$$

$$E_0 = E_0^{(0)} + \lambda E_0^{(1)} + \lambda^2 E_0^{(2)} + \cdots \tag{3.241}$$

Substitution of (3.240) and (3.241) in the Schrödinger equation for the perturbed system (3.238) leads to

$$\left[\left(\hat{H}^0 + \lambda\hat{H}'\right) - \left(E_0^{(0)} + \lambda E_0^{(1)} + \lambda^2 E_0^{(2)} \ldots\right)\right]\left(\phi_0^{(0)} + \lambda\psi_0^{(1)} + \lambda^2\psi_0^{(2)} + \cdots\right) = 0 \tag{3.242}$$

Equating the coefficients of the λ^1 terms, one obtains

$$\lambda^1: \quad \left(\hat{H}^0 - E_0^{(0)}\right)\psi_0^{(1)} + \left(\hat{H}' - E_0^{(1)}\right)\phi_0^{(0)} = 0 \tag{3.243}$$

that is,

$$\left(\hat{H}^0 - E_0^{(0)}\right)\psi_0^{(1)} + \left(\hat{H}' - E_0^{(1)}\right)\sum_{k=1}^{d} a_k\psi_{0,k}^{(0)} = 0 \tag{3.244}$$

We now multiply both members by the complex conjugate of $\psi_{0,m}^{(0)}$, where m is one of the states of the d-fold degenerate unperturbed level, and integrate over all space, obtaining

$$\sum_{k=1}^{d} a_k\left(\hat{H}'_{mk} - E_0^{(1)}\delta_{mk}\right) = 0 \tag{3.245}$$

where

$$\hat{H}'_{mk} = \left\langle\psi_{0,m}^{(0)}\middle|\hat{H}'\middle|\psi_{0,k}^{(0)}\right\rangle \tag{3.246}$$

Equation (3.245) is a set of d linear homogeneous equations in the unknowns a_k, with nontrivial solutions, provided that

$$\det\left(\hat{H}'_{mk} - E_0^{(1)}\delta_{mk}\right) = 0 \tag{3.247}$$

This is an equation of degree d in $E_0^{(1)}$, with d roots that are first-order corrections to $E_0^{(0)}$,

$$E_0^{(0)} + E_{0,1}^{(1)}, \quad E_0^{(0)} + E_{0,2}^{(1)}, \quad \ldots \quad E_0^{(0)} + E_{0,d}^{(1)} \tag{3.248}$$

If all roots of (3.247) are distinct, the degeneracy is completely removed to first order. Assuming distinct roots for (3.247), we introduce each of them in (3.245), solve this set of equations for a_2/a_1, a_3/a_1, …, a_d/a_1, and find a_1 by normalization of (3.239). For each root of (3.247), $E_{0,k}^{(1)}$, the set of coefficients determines the corresponding zeroth-order function (3.239).

3.12 Time-Dependent Perturbation Theory

3.12.1 Time-Dependent Schrödinger Equation

Before presenting the time-dependent perturbation theory, we consider an unperturbed system whose time-independent Hamiltonian is represented by $\hat{H}^0$. The time evolution of the state function Ψ^0 of this undisturbed quantum-mechanical system is given by the time-dependent Schrödinger equation,

$$-\frac{\hbar}{i}\frac{\partial \Psi^0}{\partial t} = \hat{H}^0 \Psi^0 \tag{3.249}$$

where Ψ^0 is a function of the coordinates s and time t, $\Psi^0(s, t)$. If our system is in a stationary state of energy E^0, then we can use the time-independent Schrödinger equation

$$\hat{H}^0 \Psi^0 = E^0 \Psi^0 \tag{3.250}$$

and substitute this equation in (3.249), obtaining

$$-\frac{\hbar}{i}\frac{\partial \Psi^0}{\partial t} = E^0 \Psi^0 \tag{3.251}$$

Solving this differential equation is easy, since E^0 is a constant, and we obtain

$$\Psi^0(s, t) = A(s) \exp(-iE^0 t/\hbar) \tag{3.252}$$

where $A(s) = \Psi^0(s, 0)$. Hence, $A(s)$ can be identified with the eigenfunction of the time-independent Schrödinger equation for a stationary state $\psi^0(s)$,

$$\hat{H}^0 \psi^0(s) = E^0 \psi^0(s) \tag{3.253}$$

that is, $A(s) = \psi^0(s)$. Therefore,

$$\Psi^0(s, t) = \exp(-iE^0 t/\hbar)\psi^0(s) \tag{3.254}$$

Note that

$$\Psi^{0*}\Psi^0 = \psi^{0*}\psi^0 \tag{3.255}$$

since the exponential factor of Ψ^0 in (3.254) is a phase factor $e^{i\alpha}$, and as such does not affect the probability density given by (3.255).

Let us now consider $\psi_k^0(s)$ to be eigenfunctions of the time-independent Schrödinger equation

$$\hat{H}^0\psi_k^0 = E_k^0\psi_k^0 \tag{3.256}$$

Equation (3.249) can be rewritten as

$$\left(i\hbar\partial/\partial t - \hat{H}^0\right)\Psi^0 = 0 \tag{3.257}$$

Since the operator in parentheses is linear, every linear combination of solutions of the time-dependent Schrödinger equation (3.249) such as

$$\Psi^0(s,t) = \sum_k c_k \exp(-iE_k^0 t/\hbar)\psi_k^0(s) \tag{3.258}$$

is a solution of (3.249). The time-dependence of (3.258) is given by the exponentials of imaginary exponents, which act as phase factors.

3.12.2 Time-Dependent Perturbation

We now consider a time-dependent perturbation $\hat{H}'(t)$. The time-dependent Schrödinger equation is then given by

$$-\frac{\hbar}{i}\frac{\partial\Psi}{\partial t} = \left[\hat{H}^0 + \hat{H}'(t)\right]\Psi \tag{3.259}$$

where $\Psi(s, t)$ is the new state function of the perturbed system. If we specify a value for time, for example $t = t_0$, then

$$\Psi(s,t) = \sum_k b_k(t_0)\exp(-iE_k^0 t/\hbar)\psi_k^0(s) \tag{3.260}$$

since (3.259) and (3.249) have time-independent Hamiltonians, and so the time-dependence of the state functions is expressed by phase factors [see (3.258)]. Hence, the solution of (3.259) is given by

$$\Psi(s,t) = \sum_k b_k(t)\exp(-iE_k^0 t/\hbar)\psi_k^0(s) \tag{3.261}$$

where the time-dependence of the coefficients b_k is due to the perturbation. Substitution of (3.261) in (3.259) gives

$$-\frac{\hbar}{i}\frac{\partial\left[\sum_k b_k(t)\exp(-iE_k^0 t/\hbar)\psi_k^0\right]}{\partial t} = (\hat{H}^0+\hat{H}')\left[\sum_k b_k(t)\exp(-iE_k^0 t/\hbar)\psi_k^0\right] \tag{3.262}$$

We now differentiate the first member with respect to time, use the linear behavior of $\hat{H}^0$ and $\hat{H}'$ and apply (3.256)

$$\begin{aligned}&-\frac{\hbar}{i}\sum_k \frac{db_k(t)}{dt}\exp\left(-iE_k^0 t/\hbar\right)\psi_k^0 + \sum_k b_k(t)\exp\left(-iE_k^0 t/\hbar\right)E_k^0\psi_k^0\\ &= \sum_k b_k(t)\exp\left(-iE_k^0 t/\hbar\right)E_k^0\psi_k^0 + \sum_k b_k(t)\exp\left(-iE_k^0 t/\hbar\right)\hat{H}'\psi_k^0\end{aligned}$$

Canceling equal terms on both sides of this equation gives

$$-\frac{\hbar}{i}\sum_k \frac{\mathrm{d}b_k(t)}{\mathrm{d}t}\exp\left(-iE_k^0 t/\hbar\right)\psi_k^0 = \sum_k b_k(t)\exp\left(-iE_k^0 t/\hbar\right)\hat{H}'\psi_k^0 \tag{3.263}$$

In order to select the time derivative of $b_n(t)$ on the left-hand side of this equation, we multiply both members by ψ_n^{0*}, integrate over the spatial and spin coordinates, and use

$$\left\langle\psi_n^0|\psi_k^0\right\rangle = \delta_{nk} \tag{3.264}$$

obtaining

$$\frac{\mathrm{d}b_n(t)}{\mathrm{d}t} = -\frac{i}{\hbar}\sum_k b_k(t)\exp(i\omega_{nk}t)\left\langle\psi_n^0|\hat{H}'|\psi_k^0\right\rangle$$

where

$$\omega_{nk} = \frac{E_n^0 - E_k^0}{\hbar}$$

As with time-independent perturbation theory [see (3.210)], we introduce now the scaling parameter λ, rewrite the above equation as

$$\frac{\mathrm{d}b_n(t)}{\mathrm{d}t} = -\frac{i}{\hbar}\sum_k b_k(t)\exp(i\omega_{nk}t)\left\langle\psi_n^0|\lambda\hat{H}'|\psi_k^0\right\rangle \tag{3.265}$$

and expand the coefficients $b_k(t)$ in powers of λ (see *Ballentine*, Further Reading, Sect. 12.5),

$$b_k(t) = b_k^{(0)} + \lambda\, b_k^{(1)}(t) + \lambda^2 b_k^{(2)}(t) + \cdots \tag{3.266}$$

We now substitute (3.266) in (3.265), collect equal powers of λ, and obtain, to first order,

$$\frac{\mathrm{d}b_n^{(1)}(t)}{\mathrm{d}t} = -\frac{i}{\hbar}\sum_k b_k^{(0)} \exp(i\omega_{nk}t)\left\langle \psi_n^0 \middle| \hat{H}' \middle| \psi_k^0 \right\rangle \tag{3.267}$$

and to second order,

$$\frac{\mathrm{d}b_n^{(2)}(t)}{\mathrm{d}t} = -\frac{i}{\hbar}\sum_k b_k^{(1)}(t) \exp(i\omega_{nk}t)\left\langle \psi_n^0 \middle| \hat{H}' \middle| \psi_k^0 \right\rangle$$

In (3.267), the time-independent zeroth-order coefficients correspond to the absence of perturbation and to stationary states. In addition, integration of (3.267) enables us to obtain the first-order coefficient $b_n^{(1)}(t)$.

Before the time-dependent perturbation $\hat{H}'(t)$ is applied at $t = 0$, let us assume that the system is in the nondegenerate stationary energy level E_i^0, and so $b_i^{(0)} = 1$ and $b_k^{(0)} = 0$ for $k \neq i$, that is,

$$\Psi_i^0 = \exp(-iE_i^0 t/\hbar)\psi_i^0 \tag{3.268}$$

The perturbation is switched on in the time interval $0 < t < \tau$, during which the coefficients of (3.261) become variable, and so there are no stationary states. For times $t \geq \tau$, the perturbation vanishes, and the coefficients retain the values they had at time τ, that is, $b_k(t > \tau) = b_k(\tau)$. The probability of having a nondegenerate state with energy E_f^0 is given by $|b_f(\tau)|^2$ [see (3.261)]. In order to obtain this probability, we begin by using (3.267) and write

$$\begin{aligned}\frac{\mathrm{d}b_f^{(1)}(t)}{\mathrm{d}t} &= -\frac{i}{\hbar}\sum_k b_k^{(0)} \exp(i\omega_{fi}t)\left\langle \psi_f^0 \middle| \hat{H}' \middle| \psi_k^0 \right\rangle \\ &= -\frac{i}{\hbar}\exp(i\omega_{fi}t)\left\langle \psi_f^0 \middle| \hat{H}' \middle| \psi_i^0 \right\rangle \quad (b_k^{(0)} = \delta_{ki})\end{aligned} \tag{3.269}$$

We now integrate both members between $t = 0$ and $t = \tau$, use (3.266) for $\lambda = 1$, and note that $b_f^{(0)} = 0$, obtaining

$$b_f(\tau) \approx b_f^{(1)}(\tau) = -\frac{i}{\hbar}\int_0^\tau \left\langle \psi_f^0 \middle| \hat{H}' \middle| \psi_i^0 \right\rangle \exp(i\omega_{fi}t)\mathrm{d}t \tag{3.270}$$

Finally, we take the square of the absolute value of both members to calculate the probability of having the state with energy E_f^0:

$$\left|b_f(\tau)\right|^2 \approx \frac{1}{\hbar^2}\left|\int_0^\tau \left\langle \psi_f^0 \middle| \hat{H}' \middle| \psi_i^0 \right\rangle \exp(i\omega_{fi}t)\mathrm{d}t\right|^2 \tag{3.271}$$

3.13 Absorption and Emission of Radiation

We now consider the interaction of a molecule with electromagnetic radiation, where the molecule is dealt with quantum-mechanically but the electromagnetic radiation is treated classically. A plane polarized electromagnetic radiation consists of one electric field and one magnetic field whose amplitudes are sinusoidal waves that travel in phase in mutually perpendicular planes. Electrons in atoms or molecules interact mostly with the electric field. The mathematical expression for a plane polarized electric field wave is given by

$$\mathcal{E}_x(z,t) = \mathcal{E}_0 \sin(\omega t - kz) \tag{3.272}$$

where the electric field $\mathcal{E}_0$ is the amplitude of $\mathcal{E}_x$, is polarized in the x direction, propagates along z, and

$$\omega = \frac{2\pi}{\tau} = 2\pi\nu \quad k = \frac{2\pi}{\lambda} = 2\pi\tilde{\nu} \tag{3.273}$$

where ω is the angular frequency (τ is the period of the radiation and ν its frequency), and k is the angular wavenumber (λ is the wavelength, its inverse being the wavenumber). The electric field (the electric force per unit charge) interacts with charges in the molecule, electrons, and nuclei. The force on charge q is $F = \mathcal{E}_x q = -\mathrm{d}V/\mathrm{d}x$, where V is the potential energy. Integration of the latter equation gives $V = -\mathcal{E}_x q x$. Therefore, for a distribution of charges and an x-polarized electric field, the perturbation is given by

$$\hat{H}'(t) = -\mathcal{E}_x \sum_\alpha q_\alpha x_\alpha = -\mathcal{E}_0 \sin(\omega t - kz) \sum_\alpha q_\alpha x_\alpha \tag{3.274}$$

where q_α represent the charges and x_α their positions along x. Substitution of (3.274) in (3.270) gives

$$b_f(\tau) \approx \frac{i\mathcal{E}_0}{\hbar}\int_0^\tau \left\langle \psi_f^0 \middle| \sum_\alpha q_\alpha x_\alpha \middle| \psi_i^0 \right\rangle \sin(\omega t - kz)\exp(i\omega_{fi})\mathrm{d}t \tag{3.275}$$

Note that the wavelength of ultraviolet radiation is of order 10^2 nm, and visible and infrared radiations have larger wavelengths. Hence, the term kz in the argument of the sine function is small, since k is proportional to the inverse of the wavelength [see (3.273)]. In addition, the order of magnitude of the largest dimension of a typical molecule is 1 nm, and so ψ_i^0 and ψ_f^0 are vanishingly small for distances greater than 1 nm. Hence, using the so-called **long wavelength approximation**, we can replace (3.275) by the simpler expression

$$b_f(\tau) \approx \frac{i\mathcal{E}_0}{\hbar} \int_0^\tau \left\langle \psi_f^0 \middle| \sum_\alpha q_\alpha x_\alpha \middle| \psi_i^0 \right\rangle \sin(\omega t) \exp(i\omega_{fi} t) \mathrm{d}t \tag{3.276}$$

which, in turn, can be replaced by

$$b_f(\tau) \approx \frac{i\mathcal{E}_0}{\hbar} \left\langle \psi_f^0 \middle| \sum_\alpha q_\alpha x_\alpha \middle| \psi_i^0 \right\rangle \int_0^\tau \sin(\omega t) \exp(i\omega_{fi} t) \mathrm{d}t \tag{3.277}$$

since ψ_i^0 and ψ_f^0 are time-independent wave functions and the dipole moment operator does not explicitly depend on time. Substitution of

$$\sin(\omega t) = \frac{\exp(i\omega t) - \exp(-i\omega t)}{2i} \tag{3.278}$$

in (3.277) gives

$$b_f(\tau) \approx \frac{\mathcal{E}_0}{2\hbar} \left\langle \psi_f^0 \middle| \sum_\alpha q_\alpha x_\alpha \middle| \psi_i^0 \right\rangle \int_0^\tau \left\{ \exp[i(\omega_{fi} + \omega)t] - \exp[i(\omega_{fi} - \omega)t] \right\} \mathrm{d}t \tag{3.279}$$

In turn, use of

$$\int_0^\tau \exp(at) \mathrm{d}t = \frac{\exp(a\tau) - 1}{a} \tag{3.280}$$

leads to

$$b_f(\tau) \approx \frac{\mathcal{E}_0}{2\hbar i} \left\langle \psi_f^0 \middle| \sum_\alpha q_\alpha x_\alpha \middle| \psi_i^0 \right\rangle \left\{ \frac{\exp[i(\omega_{fi} + \omega)\tau] - 1}{\omega_{fi} + \omega} - \frac{\exp[i(\omega_{fi} - \omega)\tau] - 1}{\omega_{fi} - \omega} \right\} \tag{3.281}$$

The first and second terms of the expression in braces lead to 0/0 indeterminate forms for $\omega = -\omega_{fi}$ and $\omega = \omega_{fi}$, respectively, which can be solved by taking the following limits:

$$\lim_{\omega_{fi} \pm \omega \to 0} \frac{\exp[i(\omega_{fi} \pm \omega)\tau] - 1}{\omega_{fi} \pm \omega} = i\tau \tag{3.282}$$

(**E41**). These limits can be easily obtained by the quotient of derivatives of the numerator and denominator, according to l'Hôpital's rule.

Considering the square of the absolute value of the expression in braces in (3.281), Fig. 3.11 plots the resulting band shape as a function of ω in a neighborhood of ω_{fi}. The maximum at $\omega = \omega_{fi}$ corresponds to an ascending transition from i to f that leads to **absorption** of energy from the electromagnetic radiation. In turn, the maximum at $\omega = -\omega_{fi}$ corresponds to a downward transition from f to i that leads to **stimulated emission** of electromagnetic radiation, a radiation process found in lasers.

The probability of the spectroscopic transition from i to f is given by the square of the absolute value of b_f [see (3.281)]. If the radiation has additional electric field components along y and z, then we have to add the squares of the absolute values of (3.281) for the corresponding y and z terms and separate the square of the band shape factor in (3.281) (see Fig. 3.11), being left with

$$\left|\left\langle \psi_f^0 | \hat{\mu}_x | \psi_i^0 \right\rangle\right|^2 + \left|\left\langle \psi_f^0 | \hat{\mu}_y | \psi_i^0 \right\rangle\right|^2 + \left|\left\langle \psi_f^0 | \hat{\mu}_z | \psi_i^0 \right\rangle\right|^2 \tag{3.283}$$

For the i to f transition probability to be significant, at least one of the above integrals should be different from zero. We come back to this point in Sect. 4.5 to discuss the implications of molecular symmetry.

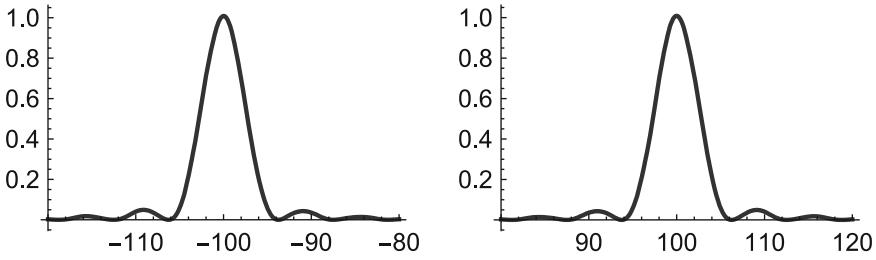


Fig. 3.11 Band shapes for emission ($\omega = -\omega_{fi}$) and absorption ($\omega = \omega_{fi}$) as functions of ω. The units are arbitrary. Figure obtained with *Mathematica*

3.13.1 Spontaneous Emission of Radiation

Both absorption and stimulated emission of radiation are induced by radiation and have transition probabilities proportional to the spatial density of radiation, d(ω). The probability of absorption per unit of radiation density and unit of population of the initial state is called for historical reasons the **Einstein *B* coefficient** for absorption and is equal to the Einstein coefficient for stimulated emission, thus being represented by the same letter. The coefficient for absorption and stimulated emission is proportional to the square of the transition dipole moment (3.283).

Let us consider two nondegenerate quantum states a and b, where b is the excited state. The equality between Einstein coefficients for absorption and stimulated emission implies the occurrence of transitions that are independent of the radiation density, thus being associated with spontaneous emission. If spontaneous emission did not exist, state b would acquire the same population as state a, and then it would not be possible to observe the transition between these states spectroscopically. The probability of **spontaneous emission** per unit of population of the initial state is called for historical reasons the **Einstein *A* coefficient**.

At equilibrium, the increase in the population of state b is balanced with the decrease of this population, and we can write

$$B\mathrm{d}\rho_a - B\mathrm{d}\rho_b - A\rho_b = 0 \tag{3.284}$$

where ρ_a and ρ_b are the populations of the a and b states, the first term is the probability for absorption, the second term stands for the probability of stimulated emission, and the last term accounts for the probability of spontaneous emission. At equilibrium and a specified temperature, ρ_a and ρ_b are related by the Boltzmann distribution, which is given by

$$\rho_b = \rho_a \exp(-\hbar\omega_{ba}/kT) \tag{3.285}$$

where $\omega_{ba} = \omega_b - \omega_a$ and k is the Boltzmann constant ($k = 1.3806488 \times 10^{-23}$ J K^{-1}; see Appendix). Substitution of (3.285) in (3.284) leads to

$$A = Bd[\exp(-\hbar\omega_{ab}/kT) - 1] \tag{3.286}$$

We now need to know the radiation density d. Using the spatial density of radiation in equilibrium with the blackbody at a defined temperature, we write, according to Planck's law,

$$d \propto \omega^3[\exp(-\hbar\omega/kT) - 1]^{-1} \tag{3.287}$$

where the symbol $\propto$ stands for "proportional to." Substitution of (3.287) in (3.286) gives

$$A \propto B\omega^3 \frac{\exp(-\hbar\omega_{ab}/kT) - 1}{\exp(-\hbar\omega/kT) - 1} \tag{3.288}$$

As mentioned before [see (3.281) and Fig. 3.11], B takes nonnegligible values in the resonance condition, that is, when $\omega = \omega_{ab}$. Hence, (3.288) can be replaced by

$$A \propto B\omega^3 \tag{3.289}$$

This result shows that spontaneous emission is increasingly important for high frequencies and justifies the difficulties in obtaining laser radiation in the ultraviolet region.

3.14 Raman Scattering

In 1928, the Indian physicist C.V. Raman (1888–1970) experimentally showed that a small fraction of a monochromatic beam of visible light is scattered by a sample with change of frequency (inelastic light scattering) and that the observed shifts in frequency depend on the chemical structure of the molecules responsible for the scattering. For this discovery, Raman was awarded the 1930 Nobel Prize in physics.

Light scattering involves incident and scattered photons. Most of the latter are scattered with the same frequency as the incident photons (elastic or Rayleigh scattering, **E42**), and a small fraction of them are scattered at a different frequency (inelastic or Raman scattering). In Rayleigh scattering, the energies of the photon and molecule are separately conserved, whereas in Raman scattering, energy conservation is observed for the whole radiation–molecule system, that is, frequency changes in the radiation and molecule are of equal magnitude but opposite signs. In Raman scattering, Stokes transitions are those in which the scattered photons lose energy to the molecule, whereas anti-Stokes transitions are those in which the scattered photons gain energy from the molecule. Figure 3.12 presents the Rayleigh and Raman spectra of carbon tetrachloride (liquid) according to the IUPAC (International Union of Pure and Applied Chemistry) recommendations, where the abscissa is a linear scale occupied by the Raman shifts presented in cm^{-1} and decreasing from left to right (the frequency of the scattered light increases from left to right, as it should), and the ordinate is linear and proportional to the intensity. The intensity of Raman scattering is several orders of magnitude lower than the intensity of Rayleigh scattering. The distribution of Stokes and corresponding anti-Stokes lines on each side of the Rayleigh line is symmetric in frequency, but Stokes lines are more intense than the corresponding anti-Stokes lines (Fig. 3.12). For this reason, a Raman spectrum usually records only the Stokes side of the spectrum.

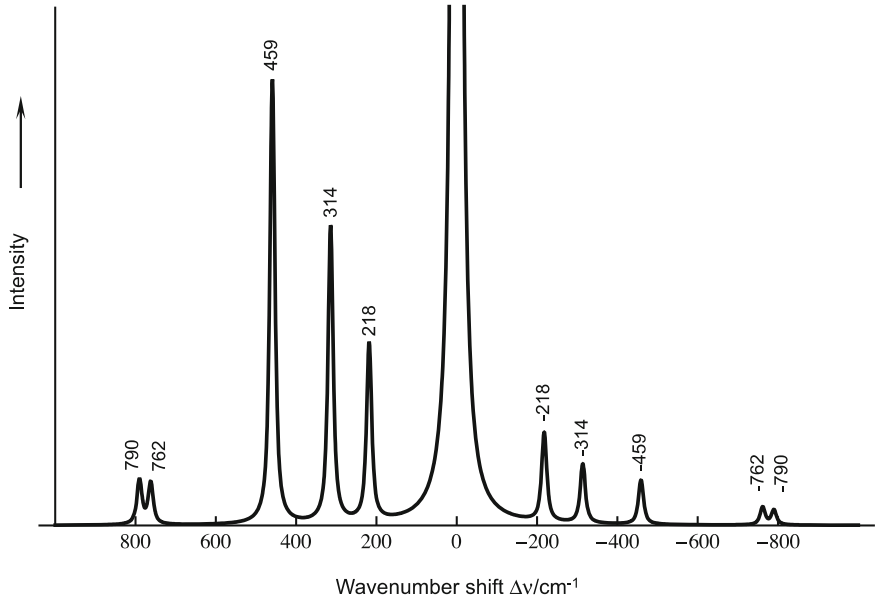


Fig. 3.12 Rayleigh and Raman spectra of carbon tetrachloride (liquid) presented according to IUPAC recommendations. Details of band shapes and band intensities should be considered schematic. This figure was obtained using *Mathematica* with each band simulated by a linear combination of Lorentzian and Gaussian functions

3.14.1 Classical Treatment

The electric field of electromagnetic radiation induces an electric dipole moment in a molecule that is given by

$$\boldsymbol{\mu}_{\text{ind}} = \boldsymbol{\alpha}\mathcal{E} \tag{3.290}$$

where the induced dipole moment and the electric field are represented by column vectors in the $(xyz)^{\text{T}}$ basis, and the polarizability tensor is represented by a 3×3 symmetric matrix

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{pmatrix} \tag{3.291}$$

(**E43**). In the following equalities of this subsection, we assume that the polarizability is spherical ($\alpha_{xx} = \alpha_{yy} = \alpha_{zz} = \alpha$ and off-diagonal elements equal to zero) to make the equations simpler without unnecessary subscripts. The electric field of a plane-polarized electromagnetic radiation varies in time with the frequency of the radiation ν_0 and is given by

$$\mathcal{E} = \mathcal{E}_0 \cos(2\pi\nu_0 t) \tag{3.292}$$

In turn, the polarizability changes with the vibrational coordinate involved in the Raman transition, and its expansion in a Taylor series gives

$$\alpha = \alpha_0 + \left(\frac{\partial\alpha}{\partial Q}\right)_0 Q + \cdots \tag{3.293}$$

where the vibrational coordinate Q has a specific frequency and is given by

$$Q = A\cos(2\pi\nu t) \tag{3.294}$$

Substitution of (3.292), (3.293), and (3.294) in (3.290) gives, to first order,

$$\mu_{\text{ind}} = \alpha_0\mathcal{E}_0\cos(2\pi\nu_0 t) + \left(\frac{\partial\alpha}{\partial Q}\right)_0 A\mathcal{E}_0\cos(2\pi\nu_0 t)\cos(2\pi\nu t) \tag{3.295}$$

Using simple trigonometry [cos $(a + b)$ = cos$a$ cos$b$ − sin$a$ sin$b$], one obtains

$$\mu_{\text{ind}} = \alpha_0\mathcal{E}_0\cos(2\pi\nu_0 t) + \frac{1}{2}\left(\frac{\partial\alpha}{\partial Q}\right)_0 A\mathcal{E}_0\{\cos[2\pi(\nu_0+\nu)t] + \cos[2\pi(\nu_0-\nu)t]\} \tag{3.296}$$

This expression shows that the frequency ν_0 corresponds to the Rayleigh scattering and involves the static polarizability α_0, whereas the Stokes and anti-Stokes Raman frequencies $\nu_0 - \nu$ and $\nu_0 + \nu$ involve the polarizability derivative with respect to the vibrational coordinate $(\partial\alpha/\partial Q)_0$ (**E44**). A Raman active vibration requires

$$\left(\frac{\partial\alpha}{\partial Q}\right)_0 \neq 0 \tag{3.297}$$

with the Raman intensity being proportional to the square of the polarizability derivative with respect to the vibrational coordinate.

3.14.2 Quantum-Mechanical Treatment

While the quantum-mechanical treatment of absorption and stimulated emission of radiation led to the electric dipole transition integrals in (3.283), the quantum-mechanical treatment of light scattering considers the first-order induced transition electric dipole moment, whose ρ component (ρ can be x, y, or z) is given by

$$\left(\mu_{\rho}^{(1)}\right)_{fi} = \sum_{\sigma=x,y,z} \left\{ \left(\alpha_{\rho\sigma}\right)_{fi} \mathcal{E}_{\sigma 0} \exp\left[-i\left(\omega_0 - \omega_{fi}\right)\right] \right\}$$

where $\varepsilon_{\sigma 0}$ is the electric field amplitude along σ (σ can be x, y, or z) (see Long, Further Reading, Sect. 4.2). Separation of the electric field [see (3.290)] in this expression leads to the transition polarizability, whose $\rho\sigma$ component between the initial and final states i and f is given by

$$\left(\alpha_{\rho\sigma}\right)_{fi} = \frac{1}{\hbar} \sum_{r \neq i,f} \left(\frac{\left\langle \psi_f^0 \middle| \hat{\mu}_{\rho} \middle| \psi_r^0 \right\rangle \left\langle \psi_r^0 \middle| \hat{\mu}_{\sigma} \middle| \psi_i^0 \right\rangle}{\omega_{ri} - \omega_0 - i\Gamma_r} + \frac{\left\langle \psi_f^0 \middle| \hat{\mu}_{\sigma} \middle| \psi_r^0 \right\rangle \left\langle \psi_r^0 \middle| \hat{\mu}_{\rho} \middle| \psi_i^0 \right\rangle}{\omega_{rf} + \omega_0 + i\Gamma_r} \right) \tag{3.298}$$

where ψ_i^0, ψ_f^0, and ψ_r^0 are time-independent unperturbed wave functions for the initial state i, the final state f, and the intermediate state r (these wave functions are assumed to be real), $\hat{\mu}_{\rho}$ is the dipole moment operator, ω_0 is the angular frequency of the incident radiation,

$$\omega_0 = E_0/\hbar \tag{3.299}$$

and

$$\omega_{ri} = \omega_r - \omega_i \tag{3.300}$$

The relation between the time-dependent and time-independent wave functions for state r is given by

$$\Psi_r = \psi_r \exp[-\mathrm{i}(\omega_r - \mathrm{i}\Gamma_r)t] \tag{3.301}$$

where Γ_r is the half-width of state r, which is related to the lifetime τ_r through the energy–time uncertainty principle, as shown below:

$$\Delta E \Delta t \sim \hbar \Rightarrow \Delta\omega\Delta t \sim 1 \Rightarrow 2\Gamma_r\tau_r \sim 1 \Rightarrow \tau_r \sim \frac{1}{2\Gamma_r} \tag{3.302}$$

Each state r is not stationary, that is to say, is not a solution of the time-independent Schrödinger equation, and so does not have a defined energy, and is therefore called a **virtual state**. In turn, the initial and final states i and f are stationary, have defined energies, i.e., $\Delta E_i = \Delta E_f = 0$ ($\Gamma_i = \Gamma_f = 0$), and so their lifetimes are infinite (the physical attributes of states i and f never change). As a result,

$$\Psi_i = \psi_i \exp(-\mathrm{i}\omega_i t) \quad \Psi_f = \psi_f \exp(-\mathrm{i}\omega_f t) \tag{3.303}$$

Note that (3.298) corresponds to a second-order energy correction of perturbation theory and to a first-order correction in the wave function [see (3.225), (3.226), and (3.227) and compare (3.298) with (3.227)].

We now consider the transition polarizability (3.298) and the cases that result from increasing the angular frequency ω_0 of the exciting radiation (see *Long*, Further Reading, Sect. 4.3). The first case is that in which ω_0 is much smaller than ω_{ri} and ω_{rf} for all states r and the half-widths Γ_r are small relative to ω_{ri} and ω_{rf} [Γ_r can be neglected in both terms of (3.298)]. Since the denominators of (3.298) then become approximately equal to ω_{ri} and ω_{rf} and the summation over r is infinite, no direct information can be obtained on the virtual states r, and this type of Raman scattering, called **normal Raman scattering** (Fig. 3.13a), is essentially a ground-state property. The states r that significantly contribute are those for which the numerators are different from zero; that is, the transition dipole moments from the state i to a virtual state r and from a state r to the final state f are both different from zero, and the weighting of these contributions is given by the inverse of the corresponding denominators, which are approximately equal to ω_{ri} and ω_{rf}, as we have mentioned above.

When ω_0 approaches one ω_{ri} transition frequency, the denominator $\omega_{ri} - \omega_0 - i\Gamma_r$ becomes appreciably smaller, due to partial cancellation in the difference $\omega_{ri} - \omega_0$, and this type of light scattering is called **preresonance Raman scattering** (Fig. 3.13b).

When $\omega_0 \approx \omega_{ri}$ for a particular state r or a limited number of states r, then $\omega_{ri} - \omega_0 - i\Gamma_r$ becomes approximately equal to $i\Gamma_r$, the corresponding states r dominate the sum in (3.298), and the Raman intensity can drastically increase; this

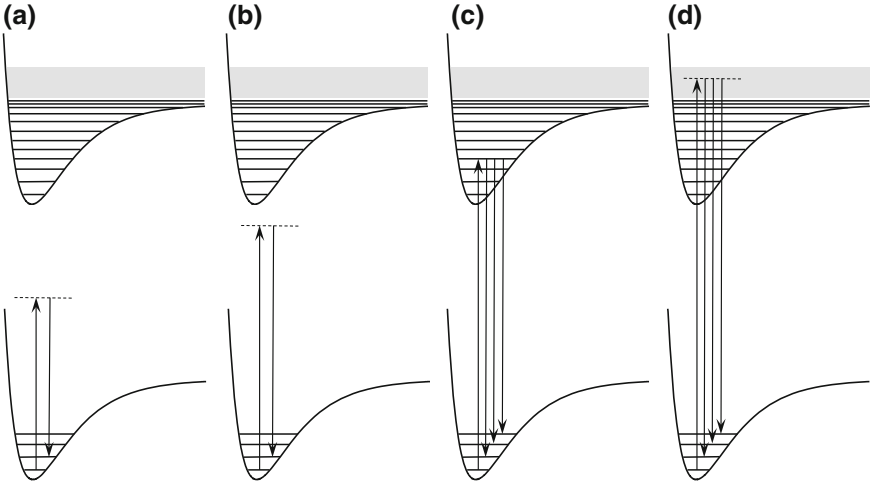


Fig. 3.13 Raman scattering processes: **a** normal Raman scattering, $\omega_0 \ll \omega_{ri}$; **b** preresonance Raman scattering, $\omega_0 \rightarrow \omega_{ri}$; **c** discrete resonance Raman scattering, $\omega_0 \approx \omega_{ri}$; **d** continuum resonance Raman scattering, ω_0 in the range of dissociative continuum levels. The potential energy curve was obtained with *Mathematica* using a Morse function to which terms in x^{-6}, x^{-8}, and x^{-10} were added

type of scattering is called **discrete resonance Raman scattering** (Fig. 3.13c). Detailed information on the selected states r and their lifetimes can then be obtained. Since the denominator $\omega_{rf} + \omega_0 + i\Gamma_r$ of (3.298) involves the sum of ω_{rf} and ω_0, cancellation does not occur, and so the corresponding terms become relatively unimportant in the sum over r.

Finally, when ω_0 falls in the range of dissociative continuum levels, we have what is called **continuum resonance Raman scattering** (Fig. 3.13d).

3.15 Molecular Calculations

Molecular calculations are usually performed within the Born–Oppenheimer approximation and comprise three major methods for solving the Schrödinger equation: ab initio, density functional theory (DFT), and semiempirical methods.

3.15.1 Computational Methods

The Schrödinger equation of **ab initio** or *first principles* calculations does not include adjustable parameters, and all integrals are evaluated exactly. The electronic Hamiltonian of these calculations includes fundamental physical constants such as the Planck constant, electron mass, and electron charge. Therefore, ab initio calculations do not lead to compensation errors, thus having a good reputation, especially in comparison with semiempirical calculations. However, in molecules with tens of atoms, they may be prohibitively time-consuming.

A method due to Moller and Plesset (1934) and applied to many-electron systems uses perturbation theory to obtain an approximation to the electron correlation energy. The unperturbed Hamiltonian is the sum of Fock operators, and the perturbation is the difference between the true interelectronic repulsions and the Hartree–Fock interelectronic potential energy. When the energy is truncated at the second-order perturbation correction, the method is called **MP2** for Moller–Plesset-2.

DFT calculations are a good alternative to ab initio calculations because they are less time-consuming and present results of comparable quality. A popular molecular DFT calculation much used in this book is known by the acronym **B3LYP**, which stands for Becke's 3-parameter exchange functional (Becke 1993) and Lee, Yang, and Parr correlation functional (Lee et al. 1988; Miehlich et al. 1989). DFT calculations with the B3LYP exchange-correlation functional generally give geometries and frequencies in close agreement with experimental results. When the number of atoms in a molecule is sufficiently large that the DFT calculation becomes too time-consuming, semiempirical calculations are a possible alternative.

In **semiempirical** calculations, molecular integrals with negligible values are taken as zero, and integrals that can be assigned some physical meaning are replaced by parameters obtained by fitting experimental data (empirical parameters). These

approximations make the calculations faster, thus becoming amenable to molecules with hundreds of atoms. Semiempirical methods are frequently used with success in organic molecules. These have a reduced number of different atoms, the most common being C and H, with O, N, P, and S occurring less frequently, and a limited number of local geometries, the more frequently encountered being the tetrahedral and planar triangular geometries. The use of empirical parameters inevitably leads to compensation of errors and to results whose agreement with experimental values is more apparent than real.

3.15.2 Gaussian-Type Functions

Orbitals of a quantum-mechanical molecular calculation are expanded in a set of basis functions, usually Gaussian-type functions or linear combinations of these, centered on the molecule's nuclei. A Cartesian Gaussian is given by the general expression

$$N_{ijk,\alpha} x^i y^i z^k \mathrm{e}^{-\alpha r^2} \tag{3.304}$$

where $N_{ijk,\alpha}$ is the normalizing constant, x, y, and z represent Cartesian coordinates with the origin at a particular nucleus, i, j, and k are nonnegative integers, r is the distance to the nucleus, and α is a positive orbital exponent. Values of the sum $i + j + k$ equal to 0, 1, and 2 correspond to ***s*-**, ***p*-** and ***d*-type Gaussians**, respectively. Each Slater-type orbital (STO) centered on a particular atom of a molecule is approximately given by a linear combination of Gaussian-type functions:

$$\phi \approx \sum_i d_i g_i \tag{3.305}$$

Linear combinations of Gaussian functions are called **contracted Gaussians**, and the individual Gaussians of each contracted Gaussian are called **primitive Gaussians**. The number of Gaussians of a particular contracted Gaussian is called the **degree of contraction**, and the coefficients are called **contraction coefficients**.

The potential energy of attraction between an electron and an atomic nucleus becomes infinite at the nucleus, since then the electron–nucleus distance is zero. As a result of this singularity, the wave function has a discontinuous first derivative at the nucleus, and the atomic orbital is said to have a **cusp** at the nucleus. This behavior contrasts with that of a Gaussian function that is continuous and has a continuous first derivative. *Mathematica* code **M9** illustrates this point and plots a 1*s* STO and an *s*-type Gaussian function for comparison (Fig. 3.14).

Consider now the normalized 1*s* radial function of a hydrogen atom,

$$R_{10}(r) = 2\mathrm{e}^{-r} \tag{3.306}$$

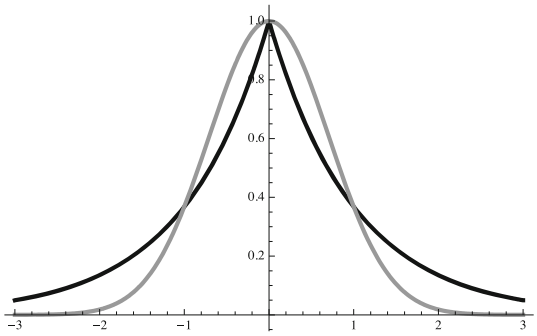


Fig. 3.14 While a 1*s* STO (*black*) has a cusp at the nucleus, the *s*-type Gaussian function (*gray*) is continuous in both the function and its derivative. These graphs were obtained with *Mathematica*

where r represents the electron–nucleus distance in bohrs (see Appendix). The angular factor for the 1*s* orbital, the spherical harmonic Y_0^0, is given by the constant $1/(2\sqrt{\pi})$. Figure 3.15 illustrates the fitting of (3.306) with a linear combination of three normalized *s*-type Gaussian functions, whose exponents were taken from the STO-3G basis set for the hydrogen atom, and the coefficients were adjustable parameters (see *Mathematica* code **M10**; **E45**).

Consider two-electron integrals of a molecular calculation. When the orbitals are expanded in a basis set of functions, the two-electron integrals may involve basis functions centered on one, two, three, or four nuclei (Fig. 3.16). In particular, integrals involving basis functions centered on three and four nuclei are very time-consuming. Gaussian functions make the calculation much easier, because the product of Gaussian functions located at two different positions with different

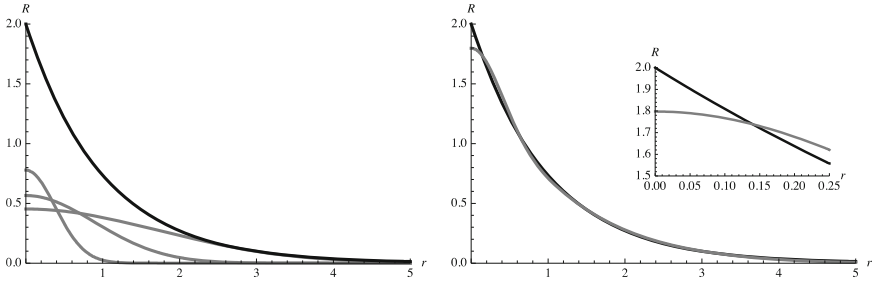


Fig. 3.15 At the *left*, the figure shows the radial factors of three *s*-type Gaussian functions and the 1*s* orbital of the hydrogen atom. At the *right*, a linear combination of the Gaussian functions is fitted to the 1*s* orbital of the hydrogen atom. The *inset* shows the cusp of the 1*s* orbital at the nucleus, contrasting with the absence of a cusp of the Gaussian function. These graphs were obtained with *Mathematica*

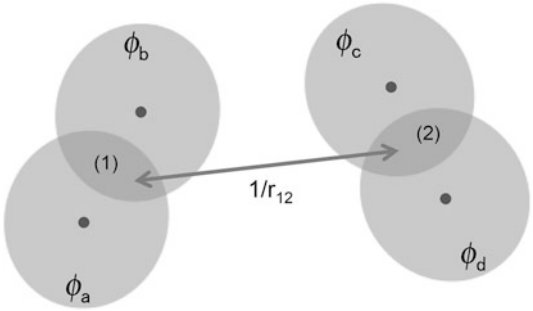


Fig. 3.16 Gaussian functions centered at four points are reduced to two Gaussian functions centered at two points

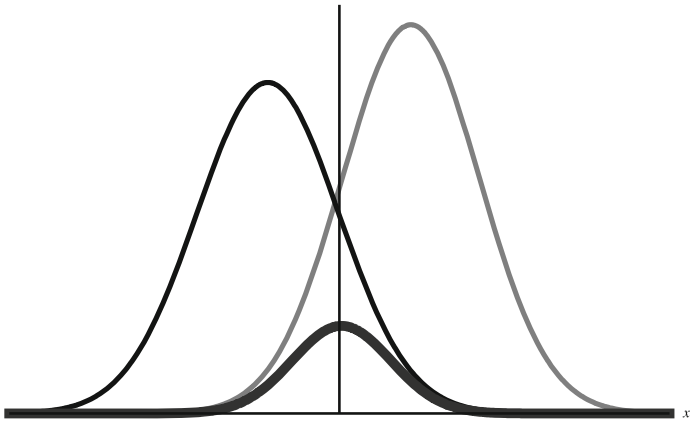


Fig. 3.17 The product of two Gaussian functions centered at different points is a third Gaussian function centered on an intermediate point. These graphs were obtained with *Mathematica*

exponents is a third Gaussian function located at an intermediate point (Fig. 3.17). To prove this statement, consider the following one-dimensional Cartesian Gaussians centered at A and B,

$$g_A(x) = N_a \mathrm{e}^{-a(x-A)^2} \quad g_B(x) = N_b \mathrm{e}^{-b(x-B)^2} \tag{3.307}$$

The product of these Gaussians gives the following Gaussian function:

$$g_C(x) = g_A(x) g_B(x) = N_a N_b \mathrm{e}^{-d} \mathrm{e}^{-c(x-C)^2} \tag{3.308}$$

where

$$d = \frac{ab}{a+b}(A-B)^2 \quad c = a+b \quad C = \frac{aA+bB}{a+b} \tag{3.309}$$

Equation (3.308) can be easily proved by showing that the sum of the exponents of g_A and g_B is equal to the exponent of g_C, that is,

$$a(x-A)^2 + b(x-B)^2 - \left[c(x-C)^2 + d\right] = 0 \tag{3.310}$$

or, after substitution of (3.309),

$$a(x-A)^2 + b(x-B)^2 - \left[(a+b)\left(x - \frac{aA+bB}{a+b}\right) + \frac{ab}{a+b}(A-B)^2\right] = 0 \tag{3.311}$$

This result is easily proved using *Mathematica* (see *Mathematica* code **M11**).

3.15.3 Standard Basis Sets

We now consider some of the more popular standard basis sets. We begin with a **minimal basis set**, i.e., a basis set with one Slater-type orbital (STO) per each inner-shell or valence-shell atomic orbital. The **STO-3G** basis set is a minimal basis set, where each STO is represented by a linear combination of three Gaussian functions. For the carbon atom, this basis set consists of five STOs (one $1s$ + one $2s$ + three $2p$ STOs) and three Gaussian functions for each STO, that is, a total of 15 Gaussian functions.

A **double-zeta basis set** consists of two STOs with distinct ζ orbital exponents for each atomic orbital, irrespective of this being an inner-shell or a valence-shell orbital. In turn, a **split-valence basis set** consists of two or more STOs per valence-shell atomic orbital, and only one STO per inner-shell atomic orbital. The **3-21G** basis set is a **valence double-zeta** basis set consisting of three primitive Gaussian functions per inner-shell atomic orbital and a split-valence shell with one contracted Gaussian function as a linear combination of two primitive Gaussian functions and an additional single primitive Gaussian function. Another example of a valence double-zeta basis set is given by the **6-31G** basis set. The basis set **6-311G** is an example of a **valence triple-zeta** basis set.

Atomic orbitals become distorted and have their centers shifted upon molecule formation. They are said to be polarized. **Polarization** is accounted for by adding STOs with ℓ quantum number greater than the maximum ℓ of the valence shell of the ground state atom. In the CH_4 molecule, the **6-31G**** basis set adds a set of five

d-type Cartesian Gaussian functions to the carbon atom (the first star) and a set of three *p*-type Cartesian Gaussian functions to each hydrogen atom (the second star). In turn, the **6-31++G**** basis set includes a set of four Gaussian functions, one *s*-type and three *p*-type functions, with very low exponents in the carbon atom (the first plus sign) and one *s*-type Gaussian function with very low exponent in each hydrogen atom (the second plus sign). These functions fall off very slowly as r increases and are called **diffuse** functions (**E46**).

Considering an *sp* valence shell, the inclusion of orbitals of higher angular momentum in the basis set should take into account the effect of each additional orbital in the lowering of the total electronic energy through its contribution to the correlation energy. For example, the lowering of energy associated with the addition of a set of *d* orbitals is approximately of the order of magnitude of one *f* orbital. In turn, the addition of a second set of *f* orbitals is equivalent to the lowering of energy caused by one *g* orbital. These facts were considered by Dunning, who proposed the so-called **correlation-consistent** (cc) basis sets (Dunning 1989). For the carbon atom, the **cc-pVTZ** (correlation-consistent polarized valence triple-zeta) basis set consists of $4s$, $3p$, $2d$, $1f$ sets of contracted Gaussian functions ($4 \times 1 + 3 \times 3 + 2 \times 5 + 1 \times 7 = 30$ functions). For the hydrogen atom, we have $3s$, $2p$, $1d$ sets of contracted Gaussian functions ($3 \times 1 + 2 \times 3 + 1 \times 5 = 14$ functions). These basis sets may be augmented with diffuse functions by adding the prefix **aug-** (for example, **aug-cc-pVTZ**).

Computing a thermodynamic energy, such as the atomization energy or the enthalpy of formation, with chemical accuracy means attaining an accuracy of 1 kcal/mol. However, the largest errors in ab initio thermodynamic calculations result from basis set truncation. Therefore, to achieve chemical accuracy, the total energy should be computed from the results of a series of calculations that extrapolate to the **complete basis set** limit.

Notes

§1. Particle in a One-Dimensional Box

A particle of mass m subjected to a potential energy that is infinite everywhere along the x-axis, except for a line segment of length ℓ where it is zero, is a particle in a one-dimensional box. The time-independent Schrödinger equation is given by

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + V(x)\psi(x) = E\psi(x) \quad 0 < x < \ell\text{: } V(x) = 0 \quad 0 \geq x \geq \ell\text{: } V(x) = \infty \tag{3.312}$$

Outside the box, that is, for $V(x) = \infty$, it is easy to conclude that the wave function is zero, i.e., the probability of finding the particle outside the box is zero, since substitution of $V(x) = \infty$ in the Schrödinger equation leads to

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = \infty\psi(x) \Rightarrow \frac{1}{\psi(x)}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = \infty \Rightarrow \psi(x) = 0 \tag{3.313}$$

Inside the box, $V(x) = 0$, and the Schrödinger equation is

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + k^2\psi(x) = 0 \tag{3.314}$$

where

$$k = \frac{(2mE)^{1/2}}{\hbar} \tag{3.315}$$

The solution for a differential equation like (3.314) is a linear combination of sine and cosine functions,

$$\psi(x) = A\cos(kx) + B\sin(kx) \tag{3.316}$$

(see **M2**), where A and B are determined by the boundary conditions. At $x = 0$, $\psi(0) = 0$, the sine function is zero, $\cos(0) = 1$, and so $A = 0$. Therefore,

$$\psi(x) = B\sin(kx) \tag{3.317}$$

The second boundary condition allows one to determine the energy, which becomes dependent on the quantum number n. In fact, at $x = l$, $\psi(l) = 0$, $B \neq 0$, and

$$\sin\left[\frac{(2mE)^{1/2}}{\hbar}\ell\right] = 0 \Rightarrow \frac{(2mE)^{1/2}}{\hbar}\ell = \pm n\pi \tag{3.318}$$

For $n = 0$, one has $E = 0$ and $k = 0$ [see (3.315)]. From (3.314), we conclude that $\psi(x)$ is zero everywhere, and so $n = 0$ is not allowed. Therefore, from the last equation we conclude that

$$E = \frac{n^2h^2}{8ml^2} \quad n = 1, 2, 3\ldots \tag{3.319}$$

Using (3.315) and (3.318), we obtain

$$k = \pm\frac{n\pi}{\ell} \tag{3.320}$$

B is the normalization constant of the wave functions, which can be easily determined using *Mathematica* (see Answer to **E8**) and gives

$$B = \left(\frac{2}{\ell}\right)^{1/2}. \tag{3.321}$$

Therefore,

$$\psi_n(x) = \left(\frac{2}{\ell}\right)^{1/2} \sin\left(\frac{n\pi x}{\ell}\right) \tag{3.322}$$

§2. Two-Particle Rigid Rotor

Consider a two-particle system with the particles held at a fixed distance d from each other. The energy of the rotor is entirely kinetic energy (rotational energy). Since the potential energy is zero, the Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla^2 \tag{3.323}$$

where

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2} \tag{3.324}$$

In spherical coordinates, the Laplacian operator takes the expression

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{1}{r^2\hbar^2}\hat{L}^2 \tag{3.325}$$

[see (3.100)]. Since the rotor is rigid, r is constant ($r = d$), and therefore we can write

$$\nabla^2 = -\frac{1}{d^2\hbar^2}\hat{L}^2 \tag{3.326}$$

Substitution of (3.326) in (3.323) leads to

$$\hat{H} = \frac{1}{2\mu d^2}\hat{L}^2 \tag{3.327}$$

Therefore, the Schrödinger equation is given by

$$\hat{H}\psi(\theta,\phi) = E\psi(\theta,\phi) \Rightarrow \frac{1}{2\mu d^2}\hat{L}^2 Y_J^m(\theta,\phi) = E_J Y_J^m(\theta,\phi) \tag{3.328}$$

[see first line of (3.91)], and the energy levels are

$$\frac{1}{2\mu d^2}J(J+1)\hbar^2 Y_J^m(\theta,\phi) = E_J Y_J^m(\theta,\phi) \quad \Rightarrow \quad E_J = \frac{J(J+1)\hbar^2}{2\mu d^2} \quad J = 0,1,2,\ldots \tag{3.329}$$

Note that the energy levels depend on J only, but the eigenfunctions depend on J and m [see (3.328) and (3.329)], where $m\hbar$ is the z-component of the rigid rotor angular momentum. This means that each specified J level has a $(2J + 1)$-degeneracy, since m takes values from $-J$ to J corresponding to different orientations of the rigid rotor angular momentum projection on a preferential direction that is quantized.

The moment of inertia I of a system of n particles about a specified axis is defined as

$$I \equiv \sum_{i=1}^{n} m_i r_i^2 \tag{3.330}$$

where m_i is the mass of the ith particle and r_i is the distance from the particle to the axis along the perpendicular to the axis. For a two-particle rigid rotor, we have

$$I = m_1 r_1^2 + m_2 r_2^2 \tag{3.331}$$

where

$$d = r_1 + r_2 \tag{3.332}$$

and r_1 and r_2 are measured from the center of mass, and so

$$m_1 r_1 = m_2 r_2 \tag{3.333}$$

From Eqs. (3.324), (3.331), (3.332), and (3.333), it can be concluded that

$$I = \mu d^2 \tag{3.334}$$

This result can be easily confirmed using *Mathematica* (see **E26**).

Mathematica Codes

M1. Wave Equation

```
sol=DSolve[∂x,xu[t,x]==(1/v^2)∂t,tu[t,x],u,{t,x}]
u[t,x]/.sol[[1]]
```

```
{{u → Function[{t, x}, C[1][-t √(v^2) + x] + C[2][t √(v^2) + x]]}}
C[1][-t √(v^2) + x] + C[2][t √(v^2) + x]
```

This *Mathematica* code considers the one-dimensional wave equation (3.18) and uses the *Mathematica* function `DSolve` for solving the differential equation for the function `u` with independent variables `t` and `x`. The symbols $\partial_{x,x}$`u[t,x]` and $\partial_{t,t}$`u[t,x]` stand for second-order partial derivatives of `u[t,x]` with respect to `x` and `t`. In general, `C[1]` and `C[2]` are default parameters generated by *Mathematica*, but in this case. they represent arbitrary functions. In the second line of code, the first level of `sol` (a list of one sublist; see the first line of results) is replaced in `u[t,x]`, thus giving the general solution of the differential equation presented in a more familiar way (see second line of results).

Suggestion: Insert the option `GeneratedParameters→f` in the first line of code. What is the effect of this change?

M2. Helmholtz Equation

```
DSolve[D[f[x],{x,2}]+k^2f[x]==0,f,x]
f[x]/.%[[1]]
```

```
{{f → Function[{x}, C[1] Cos[k x] + C[2] Sin[k x]]}}
C[1] Cos[k x] + C[2] Sin[k x]
```

This *Mathematica* code solves the so-called Helmholtz equation [see (3.20)]. The notation $\partial_{x,x}$`D[f[x],{x,2}]` represents the second derivative of `f[x]` with respect to `x`; `C[1]` and `C[2]` are default generated parameters, which in this case represent arbitrary coefficients of a linear combination of `Cos` and `Sin` functions (see second line of results).

M3. Harmonic Oscillator

Consider the one-dimensional Schrödinger equation for the harmonic oscillator and its eigenfunctions [see (3.62)], which are the product of a Gaussian function and polynomial factors known as Hermite polynomials. The built-in *Mathematica* function `HermiteH[n,z]` generates Hermite polynomials for nonnegative values of `n`. The following line of code gives the first six Hermite polynomials.

```
Table[HermiteH[n,z]],{n,0,5}]
```

$$\left\{1, 2z, -2+4z^2, -12z+8z^3, 12-48z^2+16z^4, 120z-160z^3+32z^5\right\}$$

The normalization of an eigenfunction requires the integral of the square of the eigenfunction [see (3.17)],

$$\psi_n = N_n \phi_n$$

$$\int_{-\infty}^{\infty} \psi_n^2 \mathrm{d}z = N_n^2 \int_{-\infty}^{\infty} \phi_n^2 \mathrm{d}z = 1$$

Therefore, the normalization of a harmonic oscillator eigenfunction implies the inverse of the square root of the integral of the square of the eigenfunction, as shown below:

$$N_n = \frac{1}{\sqrt{\int_{-\infty}^{\infty} \phi_n^2 \mathrm{d}z}}$$

The normalization factors are given by

$$\frac{1}{\pi^{1/4}\sqrt{2^n n!}}$$

a result that can be easily confirmed by the following *Mathematica* code:

```
aa=Table[1/Sqrt[Integrate[(Exp[-z^2/2] HermiteH[i,z])^2,{z,-Infinity,Infinity}]],{i,0,3}]
bb=Table[(1/Pi)^(1/4)/Sqrt[2^n n!],{n,0,3}]
aa-bb
```

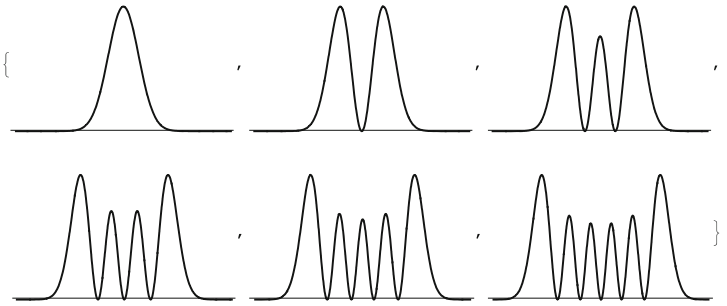
$$\left\{\frac{1}{\pi^{1/4}}, \frac{1}{\sqrt{2}\,\pi^{1/4}}, \frac{1}{2\sqrt{2}\,\pi^{1/4}}, \frac{1}{4\sqrt{3}\,\pi^{1/4}}\right\}$$

$$\left\{\frac{1}{\pi^{1/4}}, \frac{1}{\sqrt{2}\,\pi^{1/4}}, \frac{1}{2\sqrt{2}\,\pi^{1/4}}, \frac{1}{4\sqrt{3}\,\pi^{1/4}}\right\}$$

$$\{0, 0, 0, 0\}$$

The following *Mathematica* code plots the first six probability densities of the one-dimensional Schrödinger equation for the harmonic oscillator. The first line of code defines the function `f[n,z]`, and the symbol := (SetDelayed) considers the right-hand side of the expression to be the delayed value of the left-hand side. The right-hand side is evaluated each time the left-hand side appears. The result of this code is presented as a Table of graphs.

```
f[n_,z_]:=((1/Pi)^(1/4)/Sqrt[2^n n!] E^(-z^2/2) HermiteH[n,z])^2
Table[Plot[f[n,z],{z,-5,5},Axes→{True,False},Ticks→None,PlotStyle→Black],{n,0,5}]
```



Suggestion: Plot the first six Hermite polynomials in the range `{z,-5,5}`.

The following code confirms that the wave functions (3.67) are normalized and mutually orthogonal:

```
Clear[f,a]
f[n_,x_]:=(a/Pi)^(1/4)/Sqrt[2^n n!] E^(-a x^2/2) HermiteH[n,Sqrt[a]*x]
Table[Integrate[f[i,x]^2,{x,-Infinity,Infinity},Assumptions→a>0],{i,0,3}]
Table[Integrate[f[0,x]*f[i+1,x],{x,-Infinity,Infinity},Assumptions→a>0],{i,0,2}]
Table[Integrate[f[1,x]*f[i+1,x],{x,-Infinity,Infinity},Assumptions→a>0],{i,1,2}]
Integrate[f[2,x]*f[3,x],{x,-Infinity,Infinity},Assumptions→a>0]
```

```
{1, 1, 1, 1}

{0, 0, 0}

{0, 0}

0
```

M4. Spherical Harmonics

Mathematica has the built-in function `SphericalHarmonicY` for spherical harmonics. The following code lines give the mathematical expressions for spherical harmonics with $\ell = 0$, 1 and 2:

```
SphericalHarmonicY[0,0,θ,ϕ]
SphericalHarmonicY[1,{-1,0,1},θ,ϕ]
SphericalHarmonicY[2,{-2,-1,0,1,2},θ,ϕ]
```

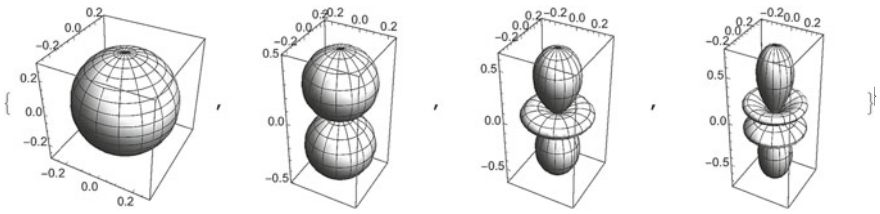
$$\frac{1}{2\sqrt{\pi}}$$

$$\left\{\frac{1}{2}e^{-i\phi}\sqrt{\frac{3}{2\pi}}\,\text{Sin}[\theta],\ \frac{1}{2}\sqrt{\frac{3}{\pi}}\,\text{Cos}[\theta],\ -\frac{1}{2}e^{i\phi}\sqrt{\frac{3}{2\pi}}\,\text{Sin}[\theta]\right\}$$

$$\left\{\frac{1}{4}e^{-2i\phi}\sqrt{\frac{15}{2\pi}}\,\text{Sin}[\theta]^2,\ \frac{1}{2}e^{-i\phi}\sqrt{\frac{15}{2\pi}}\,\text{Cos}[\theta]\,\text{Sin}[\theta],\right.$$
$$\left.\frac{1}{4}\sqrt{\frac{5}{\pi}}\,\left(-1+3\,\text{Cos}[\theta]^2\right),\ -\frac{1}{2}e^{i\phi}\sqrt{\frac{15}{2\pi}}\,\text{Cos}[\theta]\,\text{Sin}[\theta],\ \frac{1}{4}e^{2i\phi}\sqrt{\frac{15}{2\pi}}\,\text{Sin}[\theta]^2\right\}$$

The following *Mathematica* code generates 3D plots in which the absolute value of the spherical harmonic is plotted as the distance from the origin to the point, for the angles θ and ϕ. The option `Lighting→"Neutral"` specifies that the simulated lighting is a source of white light.

```
Table[SphericalPlot3D[Abs[SphericalHarmonicY[l,0,θ,ϕ]],{θ,0,Pi},
{ϕ,0,2Pi},PlotStyle→Gray,Lighting→"Neutral",ImageSize→{150,150}],{l,0,3}]
```



M5. Determinants

The following *Mathematica* code illustrates the properties of determinantal functions by confirming (3.118), (3.119), and (3.120) using 3×3 matrices and their determinants:

```
m1 = {{a11, a12, a13}, {a21, a22, a23}, {a31, a32, a33}};
Print[Matrix m1]
m1 // MatrixForm
Print[Det[m1]]
Det[m1]
m2 = {{a21, a22, a23}, {a11, a12, a13}, {a31, a32, a33}};
Print[Matrix m2: exchange of rows in m1]
m2 // MatrixForm
Print[Det[m1] / Det[m2]]
(Det[m1]/Det[m2]) // Simplify
m3 = {{(a11 + a13)/Sqrt[2], a12, (a11 - a13)/Sqrt[2]},
  {(a21 + a23)/Sqrt[2], a22, (a21 - a23)/Sqrt[2]},
  {(a31 + a33)/Sqrt[2], a32, (a31 - a33)/Sqrt[2]}};
Print[Matrix m3]
m3 // MatrixForm
Print[Det[m3] / Det[m1]]
(Det[m3]/Det[m1]) // Simplify
m4 = {{(a11 - a13)/Sqrt[2], a12, (a11 + a13)/Sqrt[2]},
  {(a21 - a23)/Sqrt[2], a22, (a21 + a23)/Sqrt[2]},
  {(a31 - a33)/Sqrt[2], a32, (a31 + a33)/Sqrt[2]}};
Print[Matrix m4]
m4 // MatrixForm
Print[Det[m4] / Det[m1]]
(Det[m4]/Det[m1]) // Simplify
m5 = {{a11, a12, a13}, {c*a11, c*a12, c*a13}, {a31, a32, a33}};
Print[Matrix m5]
m5 // MatrixForm
Print[Det[m5]]
Det[m5] // Simplify
```

```
Matrix m1
```

$$\begin{pmatrix} a11 & a12 & a13 \\ a21 & a22 & a23 \\ a31 & a32 & a33 \end{pmatrix}$$

```
Det[m1]
```

$$-a13\,a22\,a31 + a12\,a23\,a31 + a13\,a21\,a32 - a11\,a23\,a32 - a12\,a21\,a33 + a11\,a22\,a33$$

```
Matrix m2: exchange of rows in m1
```

$$\begin{pmatrix} a21 & a22 & a23 \\ a11 & a12 & a13 \\ a31 & a32 & a33 \end{pmatrix}$$

```
Det[m1]/Det[m2]
-1
Matrix m3
```

$$\begin{pmatrix} \frac{a11+a13}{\sqrt{2}} & a12 & \frac{a11-a13}{\sqrt{2}} \\ \frac{a21+a23}{\sqrt{2}} & a22 & \frac{a21-a23}{\sqrt{2}} \\ \frac{a31+a33}{\sqrt{2}} & a32 & \frac{a31-a33}{\sqrt{2}} \end{pmatrix}$$

```
Det[m3]/Det[m1]
-1
Matrix m4
```

$$\begin{pmatrix} \frac{a11-a13}{\sqrt{2}} & a12 & \frac{a11+a13}{\sqrt{2}} \\ \frac{a21-a23}{\sqrt{2}} & a22 & \frac{a21+a23}{\sqrt{2}} \\ \frac{a31-a33}{\sqrt{2}} & a32 & \frac{a31+a33}{\sqrt{2}} \end{pmatrix}$$

```
Det[m4]/Det[m1]
1
Matrix m5
```

$$\begin{pmatrix} a11 & a12 & a13 \\ a11\,c & a12\,c & a13\,c \\ a31 & a32 & a33 \end{pmatrix}$$

```
Det[m5]
0
```

M6. Systems of Homogeneous Linear Equations

The following *Mathematica* code illustrates the process of solving systems of homogeneous linear equations using 2×2 and 3×3 matrices [see (3.134), (3.136), (3.137), and (3.138)]. In the 2×2 matrix, a11 corresponds to H_{11}-$W^{(1)}S_{11}$, a12 corresponds to H_{12}-$W^{(1)}S_{12}$, a21 corresponds to H_{12}-$W^{(1)}S_{12}$, a22 corresponds to H_{22}-$W^{(1)}S_{22}$, and the determinant of the 2×2 matrix is given by (3.138):

```
Print["2x2 matrix"]
mm={{a11,a12},{a21,a22}};
mm//MatrixForm
Print["Determinant of the coefficients matrix = 0"]
Det[mm]==0
Print["Ratio r2 = c2/c1"]
Solve[a11+a12*r2==0,r2][[1]][[1]]
Solve[a21+a22*r2==0,r2][[1]][[1]]
Print["Add normalization equation and solve for c1, c2"]
Solve[c2/c1==-a11/a12&&c1^2+c2^2==1,{c1,c2}]
Print["3x3 matrix"]
nn={{a11,a12,a13},{a21,a22,a23},{a31,a32,a33}};
nn//MatrixForm
Print["Determinant of the coefficients matrix = 0"]
Det[nn] == 0
Print["Ratios r2 = c2/c1 (= A), r3 = c3/c1 (= B)"]
Solve[a11+a12*r2+a13*r3==0&&a21+a22*r2+a23*r3==0,{r2,r3}][[1]]
Print["Add normalization equation and solve for c1, c2, c3"]
Solve[c2/c1==A&&c3/c1==B&&c1^2+c2^2+c3^2==1,{c1,c2,c3}]
```

```
2x2 matrix
```

$$\begin{pmatrix} a11 & a12 \\ a21 & a22 \end{pmatrix}$$

```
Determinant of the coefficients matrix = 0
```

$$-a12\,a21 + a11\,a22 == 0$$

```
Ratio r2 = c2/c1
```

$$r2 \to -\frac{a11}{a12}$$

$$r2 \to -\frac{a21}{a22}$$

```
Add normalization equation and solve for c1, c2
```

$$\left\{\left\{c1 \to -\frac{a12}{\sqrt{a11^2 + a12^2}},\ c2 \to \frac{a11}{\sqrt{a11^2 + a12^2}}\right\},\ \left\{c1 \to \frac{a12}{\sqrt{a11^2 + a12^2}},\ c2 \to -\frac{a11}{\sqrt{a11^2 + a12^2}}\right\}\right\}$$

```
3x3 matrix
```

$$\begin{pmatrix} a11 & a12 & a13 \\ a21 & a22 & a23 \\ a31 & a32 & a33 \end{pmatrix}$$

```
Determinant of the coefficients matrix = 0
```

$$-a13\,a22\,a31 + a12\,a23\,a31 + a13\,a21\,a32 - a11\,a23\,a32 - a12\,a21\,a33 + a11\,a22\,a33 == 0$$

```
Ratios r2 = c2/c1 (= A), r3 = c3/c1 (= B)
```

$$\left\{r2 \to -\frac{a13\,a21 - a11\,a23}{a13\,a22 - a12\,a23},\ r3 \to -\frac{-a12\,a21 + a11\,a22}{a13\,a22 - a12\,a23}\right\}$$

```
Add normalization equation and solve for c1, c2, c3
```

$$\left\{\left\{c1 \to -\frac{1}{\sqrt{1 + A^2 + B^2}},\ c2 \to -\frac{A}{\sqrt{1 + A^2 + B^2}},\ c3 \to -\frac{B}{\sqrt{1 + A^2 + B^2}}\right\},\right.$$
$$\left.\left\{c1 \to \frac{1}{\sqrt{1 + A^2 + B^2}},\ c2 \to \frac{A}{\sqrt{1 + A^2 + B^2}},\ c3 \to \frac{B}{\sqrt{1 + A^2 + B^2}}\right\}\right\}$$

M7. Normalization Constants for Slater-Type Orbitals

The following *Mathematica* code computes the normalization constants of the radial factors of Slater-type orbitals for $n \leq 3$ and confirms the general expression shown in (3.158) for $n \leq 3$.

```
f[r_]:=Exp[-ζr]*r^(n-1)
norm=Table[1./Sqrt[Integrate[f[r]^2 r^2,{r,0.,Infinity},Assumptions→ζ>0]],{n,3}]
Table[Integrate[norm[[n]]^2*f[r]^2 r^2,{r,0.,Infinity},Assumptions→ζ>0],{n,3}]
Table[(2.ζ)^(n+1/2)/Sqrt[(2n)!],{n,3}]
```

$$\left\{\frac{2.}{\sqrt{\frac{1}{\zeta^3}}},\ \frac{1.1547}{\sqrt{\frac{1}{\zeta^5}}},\ \frac{0.421637}{\sqrt{\frac{1}{\zeta^7}}}\right\}$$

$$\{1., 1., 1.\}$$

$$\{2.\,\zeta^{3/2},\ 1.1547\,\zeta^{5/2},\ 0.421637\,\zeta^{7/2}\}$$

M8. Functional Derivative

Mathematica provides a built-in function for evaluating functional derivatives. It is named `VariationalD` and requires `Needs["VariationalMethods`"]`. This instruction loads an otherwise not included file in the available packages of a *Mathematica* session that is necessary for evaluating functional derivatives. The following *Mathematica* code illustrates the evaluation of functional derivatives [see (3.188) (the equation for obtaining the functional derivative) and (3.189)]:

```
Needs["VariationalMethods`"]
VariationalD[(-Z/r)*ρ[r],ρ[r],r]
VariationalD[ρ[r]*ρ[r2]/r12,ρ[r],r]
VariationalD[v[r]*ρ[r],ρ[r],r]
```

$$-\frac{Z}{r}$$

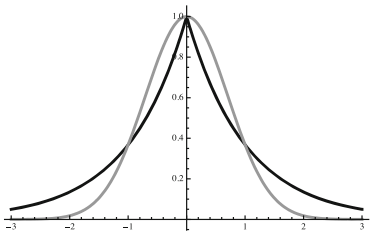
$$\frac{\rho[r2]}{r12}$$

$$v[r]$$

M9. STO Versus a Gaussian Function at the Origin

The following *Mathematica* code compares the derivatives at the origin of a 1*s* Slater-type orbital (STO) and an *s*-type Gaussian function. Since the derivative of the STO is not a continuous function at the origin, one needs to evaluate the left and right derivatives. The *Mathematica* function `Limit` finds the limiting value of a function, in this case, the ratio of differences in the dependent and independent variables, as the difference in the independent variable tends to zero. The `Limit` function has the option `Direction`, which takes the values 1 and −1 corresponding to increasing and decreasing approaches to the limit, respectively:

```
Clear[f,g]
f[x_]:=Exp[-Abs[x]]
g[x_]:=Exp[-x^2]
Plot[{f[x],g[x]},{x,-3.,3.},PlotStyle→{{Black,Thickness[0.008]},{Gray,Thickness[0.008]}},
    Ticks→None]
Limit[(f[0+h]-f[0])/h,h→0,Direction→1] (* left derivative of f[x] at x = 0- *)
Limit[(f[0+h]-f[0])/h,h→0,Direction→-1](* right derivative of f[x] at x = 0+ *)
g'[0] (* derivative of g[x] at x=0 *)
```




```
 1
-1
 0
```

M10. Fitting Gaussian Functions to a 1*s* Hydrogen Orbital

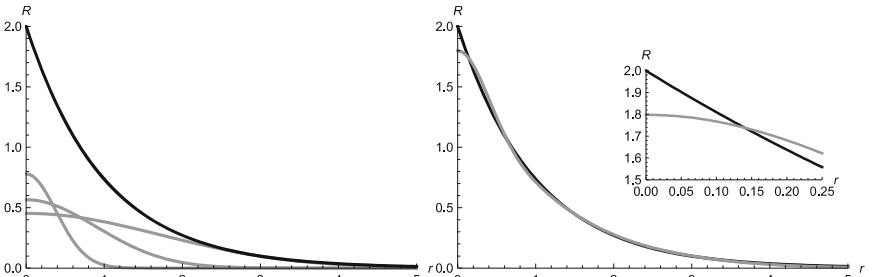
The following *Mathematica* code begins by obtaining normalization factors for the 1*s* function of the hydrogen atom and three *s*-type Gaussian functions. Being *s* functions, they depend only on the *r* coordinate, not on the angular spherical coordinates θ and ϕ, and consequently, integration is over $R^2(r)r^2\mathrm{d}r$ [see (3.105)]. After normalization, the code confirms the normalizing constants. In order to find numerical values for the Gaussians coefficients that give the best fit to `data`, the code prepares a list of lists called `data` with regularly spaced points from the function `f[r]` and defines the variable `model`. `FindFit` finds numerical values for the coefficients `c1, c2`, and `c3` that lead to the best fit of `model` to `data`. In the line of code

$$\texttt{c1} = \texttt{c1/.fit[[1]]}$$

`fit[[1]]` stands for the first part of `fit`, which is `c1→0.122444`, `c1/.fit[[1]]` replaces `c1` by `0.122444`, and `c1=c1/.fit[[1]]` assigns `0.122444` to `c1`. Similar lines of code are repeated for `c2` and `c3`. The resulting values of `c1`, `c2`, and `c3` are presented as output. The smallest coefficient (`c1`) corresponds to the Gaussian function with the highest exponent, that is, the function that does more effective fitting near the nucleus:

```
Clear[c1, c2, c3]
a = 3.425; b = 0.624; c = 0.169;
n0 = 1./Sqrt[Integrate[r^2 4*Exp[-2 r], {r, 0., Infinity}]];
n1 = 1./Sqrt[Integrate[r^2 Exp[-2 a r^2], {r, 0., Infinity}]];
n2 = 1./Sqrt[Integrate[r^2 Exp[-2 b r^2], {r, 0., Infinity}]];
n3 = 1./Sqrt[Integrate[r^2 Exp[-2 c r^2], {r, 0., Infinity}]];
f[r_] := n0*2 Exp[-r]
g1[r_] := n1*Exp[-a r^2]
g2[r_] := n2*Exp[-b r^2]
g3[r_] := n3*Exp[-c r^2]
Integrate[r^2 f[r]^2, {r, 0., Infinity}]
Integrate[r^2 g1[r]^2, {r, 0., Infinity}]
Integrate[r^2 g2[r]^2, {r, 0., Infinity}]
Integrate[r^2 g3[r]^2, {r, 0., Infinity}]
data = Table[{i/100., f[i/100.]}, {i, 0, 500}];
model = c1 g1[r] + c2 g2[r] + c3 g3[r];
fit = FindFit[data, model, {c1, c2, c3}, r];
c1 = c1 /. fit[[1]]
c2 = c2 /. fit[[2]]
c3 = c3 /. fit[[3]]
gaussians = Plot[{c1*g1[r], c2*g2[r], c3*g3[r]}, {r, 0., 5.},
    PlotRange → {{0., 5.}, {0.0, 2.0}}, PlotStyle → {{Gray, Thickness[0.008]},
    {Gray, Thickness[0.008]}, {Gray, Thickness[0.008]}},
    AxesLabel → {r, R},ImageSize → {300, 300}];
hydrogen1s = Plot[f[r], {r, 0., 5.}, PlotRange → {{0., 5.}, {0.0, 2.0}},
    PlotStyle → {Black, Thickness[0.008]}];
curves = Show[gaussians, hydrogen1s];
inset = Plot[{f[r], c1*g1[r] + c2*g2[r] + c3*g3[r]}, {r, 0., 0.25},
    PlotRange→{{0., 0.25},{1.5, 2.0}},PlotStyle→{{Black,Thickness[0.015]},
    {Gray, Thickness[0.015]}}, AxesLabel → {r, R}, ImageSize → {150, 150}];
Sum = Plot[{f[r], c1*g1[r] + c2*g2[r] + c3*g3[r]}, {r, 0., 5.},
    PlotRange→{{0.,5.},{0.,2.0}}, PlotStyle → {{Black, Thickness[0.008]},
    {Gray, Thickness[0.008]}}, AxesLabel → {r, R}, ImageSize → {300, 300},
    Epilog → Inset[inset, {3.5, 1.2}]];
Row[{curves, sum}, " "]
```

```
1.

1.

1.

1.

0.122444

0.318893

0.68047
```



M11. Product of Gaussian Functions

The following line of *Mathematica* code shows that the product of Gaussian functions located at two different positions with different exponents g_A and g_B is a third Gaussian function g_C located at an intermediate point [see (3.308)], by demonstrating that the sum of the exponents of g_A and g_B is equal to the exponent of g_C [see (3.311)]. This code illustrates the use of `FullSimplify`, which simplifies its argument (the expression to which it applies), and of `TrueQ`, which yields the result `True` or `False` for the expression to which it applies:

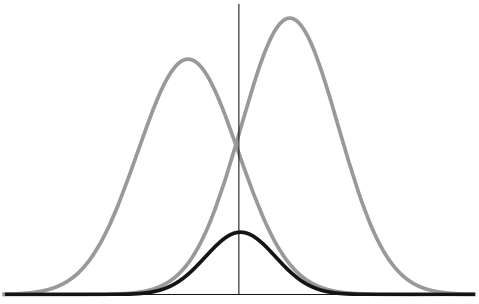
```
TrueQ[FullSimplify[a*(x-A)^2+b*(x-B)^2-((a+b)(x-(a*A+b*B)/(a+b))^2+a*b/(a+b)*(A-B)^2)]==0]
True
```

Another way of arriving at the same conclusion consists in substituting (3.309) in the left-hand side of (3.310), `FullSimplify` the result, and obtaining zero.

```
FullSimplify[a*(x-A)^2+b*(x-B)^2-(c(x-C)^2+d)/.{c→a+b,C→(a*A+b*B)/(a+b),
    d→a*b/(a+b)*(A-B)^2}]
0
```

The following *Mathematica* code illustrates the product of two Gaussians located at different positions and with different exponents by plotting the Gaussian factors and the Gaussian result:

```
x0=0.65; α=1.29; β=1.212; a=0.761; b=0.648;
g1[x_]:=a*Exp[-α (x-x0)^2]
g2[x_]:=b*Exp[-β (x+x0)^2]
Plot[{g1[x],g2[x],g1[x]*g2[x]},{x,-3.,3.},PlotRange→{{-3.,3.},{0.0,0.8}},
    PlotStyle→{{Gray,Thickness[0.008]},{Gray,Thickness[0.008]},
    {Black,Thickness[0.008]}},ImageSize→{400,400},Ticks→None]
```



Glossary

Ab initio calculations Do not include adjustable parameters, and all molecular integrals are evaluated exactly. The electronic Hamiltonian of these calculations includes fundamental physical constants such as the Planck constant, electron mass, and electron charge. Therefore, these calculations do not lead to compensation errors, as semiempirical calculations do.

Anharmonic Applied to an oscillating system that is not undergoing simple harmonic motion. The Taylor expansion of the Morse potential energy function about the equilibrium point R_e (the minimum) gives a term involving the square of $(R - R_e)$ that represents the harmonic term and cubic, quartic, and following powers of $(R - R_e)$ that account for the Morse potential anharmonicity; see (3.71).

Antisymmetry principle States that the wave function changes sign when the coordinates of two electrons including the spin are interchanged; see (3.110) for a wave function of two electrons.

Born–Oppenheimer approximation Deals with the separation between electron and nuclear coordinates in the Schrödinger equation in two steps. In the first step, the kinetic energy of the nuclei is assumed to be zero, and the Schrödinger equation is solved for the electronic Hamiltonian; see (3.140). The second step solves the Schrödinger equation for the nuclear motion; see (3.142).

Commutator Is defined by $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$. The order of application of two commuting operators is irrelevant (their commutator is equal to the null operator).

Correlation energy Given by the difference between the exact nonrelativistic energy and the Hartree–Fock energy of a molecule. A Hartree–Fock SCF wave function takes the interactions between electrons in an average way (each electron "feels" the average repulsion of all the other electrons), whereas interelectronic repulsions are instantaneous in reality, and the motions of electrons in a molecule are correlated (interdependent).

Density functional theory Based on a theorem of Pierre Hohenberg and Walter Kohn published in 1964 that says that the electron probability density of a nondegenerate ground state *uniquely* determines the Hamiltonian operator, the ground state molecular energy, and all other ground state molecular properties.

Dirac notation A notation introduced by Dirac to simplify the representation of molecular integrals; see (3.43).

Functional A rule that associates a number with a function; see (3.178) for an example of a functional.

Functional derivative For a functional J depending on the function $y(x)$ and an increment $h(x)$ on the function $y(x)$, the functional derivative of $J[y]$ with respect to y, $\delta J/\delta y$, is the limit of the quotient of $J[y + h] - J[y]$ and the area $\Delta\sigma$ lying between the curve $h(x)$ and the x-axis for an increment Δx, $(J[y + h] - J[y])/\Delta\sigma$, as $\Delta\sigma \rightarrow 0$. The functional derivative is given by (3.188).

Harmonic oscillator A single point mass m connected to a rigid wall by a perfectly elastic spring. When the perfectly elastic spring is distended or contracted, the force that restores equilibrium is proportional to the shift from the equilibrium position, $F = -kx$; see (3.52).

Hartree–Fock A method for calculating atomic and molecular energies whereby for a closed-shell electronic configuration, the wave function is a single Slater determinant of spin-orbitals that minimizes the total energy. Each electron is subject to the average Coulomb repulsion of all the other electrons, to which is added the exchange potential, which results from the antisymmetry requirement of the wave function.

Hermitian operator A linear operator that satisfies (3.27) or (3.28) for all well-behaved functions. The eigenvalues of a Hermitian operator are real numbers.

Morse potential The empirical representation of the potential energy function for a diatomic molecule given by $U(R) = D\{1 - \exp[-\beta(R - R_e)]\}^2$; see (3.68).

Reduced mass For a diatomic oscillator with masses m_A and m_B, its inverse is equal to the sum of the inverses of the masses m_A and m_B; see (3.51).

Semiempirical calculations Have molecular integrals with negligible values taken as zero and integrals that can be assigned some physical meaning replaced by parameters obtained by fitting experimental data (empirical parameters). These approximations make calculations faster and amenable to molecules with hundreds of atoms. The use of empirical parameters inevitably leads to compensation of errors and to results whose agreement with experimental values is more apparent than real.

Spherical harmonics Are eigenfunctions of the square and z-component of the orbital angular momentum operators; see (3.91) and (3.92). The square of the angular momentum operator is a function of θ and ϕ, and the z-component of the angular momentum is a function of ϕ; see (3.89) and (3.90).

Spontaneous emission of radiation Independent of radiation density. The probability of spontaneous emission per unit of population of the initial state is usually called the Einstein coefficient A. Spontaneous emission is proportional to the cube of frequency [see (3.289)], thus becoming increasingly important for high frequencies.

Stimulated emission of radiation Induced by radiation. Both absorption and stimulated emission of radiation are induced by radiation and have transition probabilities proportional to the spatial density of radiation. The probability of absorption per unit of radiation density and unit of population of the initial state is usually called the Einstein coefficient B for absorption and is equal to the Einstein coefficient for stimulated emission, thus being represented by the same letter.

Exercises

E1. Find the eigenvalues and eigenvectors of the matrix $A = \begin{pmatrix} -1 & 1 \\ 4 & 2 \end{pmatrix}$.

E2. Use *Mathematica* to find the eigenvalues and eigenvectors of the matrix $A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix}$.

E3. Show that $\partial/\partial x$ is a linear operator.

E4. Calculate the commutator between $\partial/\partial x$ and x.

E5. Relate the kinetic energy operator of the Hamiltonian to the linear momentum operator.

E6. Let u be any twice differentiable function. Show that $u(x - vt)$ is a solution of the wave equation and a forward propagating wave with speed v.

E7. Confirm that $A\cos(kx - \omega t)$ is a solution of the wave equation.

E8. Consider a particle in a one-dimensional box. Use *Mathematica* to determine the constant B for the wave functions $\psi_n(x) = B\sin(n\pi x/\ell)$.

E9. Use *Mathematica* to plot the wave functions for the particle in a one-dimensional box for $n = 1, 2, 3$.

E10. Use *Mathematica* to solve the equation $y''(x) + y'(x) - 6y(x) = 0$.

E11. Beginning with (3.27) for the definition of a Hermitian operator, arrive at (3.28). [*Hint*: Consider ψ as the linear combination $f + cg$, where c is an arbitrary complex number.]

E12. Derive (3.50) for a diatomic molecule AB.

E13. Write the equation to obtain the reduced mass of a system of four masses, m_A, m_B, m_C, and m_D.

E14. Show that (3.54) is a solution of the differential Eq. (3.53) of the classical harmonic oscillator.

E15. Use *Mathematica* to list the first six Hermite polynomials.

E16. Use *Mathematica* to find the normalization constants of the first six harmonic oscillator eigenfunctions assuming that α is equal to 1 and confirm that the normalization constants satisfy the formula $[\pi^{1/4}\sqrt{(2^n n!)}]^{-1}$.

E17. Confirm the recurrence equality (3.64) for $H_4(z)$, $H_3(z)$, and $H_2(z)$.

E18. Confirm (3.70).

E19. Derive (3.83).

E20. Derive the first of the equalities (3.86).

E21. Derive the first of the equalities (3.88).

E22. Use *Mathematica* to obtain polar plots of $|\sin(n\theta)|$, where n takes integer values from 1 to 4 and θ ranges from 0 to π.

E23. Use *Mathematica* to write an expression for $Y_2^0(\theta, 0)$ and make a polar plot of this function.

E24. Derive (3.101).

E25. Write R_{2s} in SI units.
E26. Consider the two-particle rigid rotor and the definition of moment of inertia given by $I = m_1 r_1^2 + m_2 r_2^2$. Using *Mathematica*, verify $I = \mu d^2$, where $d = r_1 + r_2$, μ is the reduced mass, and $m_1 r_1 = m_2 r_2$.
E27. Derive (3.134).
E28. Use *Mathematica* to solve the system of inhomogeneous linear equations $x + y - z = 2$, $2x - y + 3z = 5$, $3x + 2y - 2z = 5$.
E29. Use *Mathematica* to multiply two 3×3 matrices.
E30. Show that the equations of the system of homogeneous linear equations (3.137) give the same ratio of coefficients c_2/c_1.
E31. Write the equivalent of (3.147) for an atom with n electrons.
E32. Why is the term V_{NN} included in (3.154)?
E33. Write the Hartree–Fock energy expression for the ground-state $1s^2 2s^2$ electron configuration of the beryllium atom.
E34. Show that exchange integrals occur only between spin-orbitals of the same spin.
E35. Use Dirac notation to express (3.155) and (3.156).
E36. Use *Mathematica* to obtain the normalization constant for the radial factor of a Slater-type orbital with $n = 2$ and verify the result.
E37. Write (3.157) in SI units.
E38. Confirm (3.197), using the wave function (3.122).
E39. Confirm (3.199), using the wave function (3.122).
E40. Consider a hydrogen atom in its ground state perturbed by a uniform electric field applied in the z direction and represent the excited state by a $2p_z$ Slater-type orbital. Use *Mathematica* to minimize and plot the variational perturbation upper bound to the second-order energy correction as a function of the $\zeta 2p_z$ exponent.
E41. Confirm (3.282).
E42. The intensity of Rayleigh scattering is proportional to the inverse of the fourth power of wavelength. Explain why the sky is blue.
E43. Use *Mathematica* to show that the trace of the polarizability tensor and its anisotropy, whose square is given by $\gamma^2 = [(\alpha_{xx} - \alpha_{yy})^2 + (\alpha_{yy} - \alpha_{zz})^2 + (\alpha_{zz} - \alpha_{xx})^2 + 6\ (\alpha_{xy}^2 + \alpha_{yz}^2 + \alpha_{zx}^2)]/2$, are invariants under an orthogonal transformation due to a rotation of x and y around z, by an angle θ.
E44. Find the ratio of intensities of a Stokes line over the corresponding anti-Stokes line.
E45. Use *Mathematica* to plot the function $e^{-c|x|}$ with interactive manipulation of the parameter c between 0.7 and 2.0, in steps of 0.1. Denote c by ζ and present the function for $\zeta = 1.5$.
E46. Using the basis set 6-311G*, evaluate the energy difference $E_{MP2} - E_{HF}$ in kcal mol^{-1} for the C_2H_6 molecule and the chemical reaction $C_2H_4(g) + H_2(g) \rightarrow C_2H_6(g)$, at 298 K (*Suggestion*: go to CCCBDB). Comment on the obtained energy differences.

References

Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98:5648–5652
Coulson CA, Neilson AH (1961) Proc. Phys. Soc. London 78:831–837
Dunning TH Jr (1989) Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J Chem Phys 90:1007–1023
Hohenberg P, Kohn W (1964) Inhomogeneous electron gas. Phys Rev 136:B864–B871
Kohn W, Sham LJ (1965) Self-consistent equations including exchange and correlation effects. Phys Rev 140:A1133–A1138
Lee C, Yang W, Parr RG (1988) Development of the Colle-Salveti correlation-energy formula into a functional of the electron density. Phys Rev B 37:785–789
Miehlich B, Savin A, Stoll H, Preuss H (1989) Results obtained with the correlation-energy density functionals of Becke, and Lee, Yang and Parr. Chem Phys Lett 157:200–206
Moller C, Plesset MS (1934) Note on an approximation treatment for many-electron systems. Phys Rev 46:618–622

Further Reading

Ballentine LE (1998) Quantum mechanics: a modern development. World Scientific Publishing Co. Pte. Ltd.
Gelfand IM, Fomin SV (2000) Calculus of variations. Dover Publications, Inc.
Hameka HF (2004) Quantum mechanics: a conceptual approach. Wiley
Koch W, Holthausen MC (2001) A chemist's guide to density functional theory, 2nd edn. Wiley-VCH
Levine IN (2013) Quantum chemistry, 7th ed. Pearson Education Inc.
Long DA (2002) The Raman effect: a unified treatment of the theory of Raman scattering by molecules. Wiley
Merzbacher E (1998) Quantum mechanics. 3rd ed. Wiley
Parr RG, Yang W (1989) Density-functional theory of atoms and molecules. Oxford University Press
Schatz GC, Ratner MA (1993) Quantum mechanics in chemistry. Prentice-Hall International

Molecular Symmetry

4

Abstract

Symmetry manifests itself in the shapes of things, from macroscopic objects to molecules and atoms. Within the Born–Oppenheimer approximation, the nuclear framework of a molecule at equilibrium defines its symmetry, which reflects on molecular orbitals and molecular vibrations. This chapter begins by defining symmetry operations and presenting the symmetry elements of several molecules. We show that the set of symmetry operations of a molecule form a group and present a dichotomy diagram for finding the point group in a systematic way. We then represent reflection in a plane, rotation about an axis, and inversion in a point by 3×3 matrices in the (x, y, z)-basis, using these to illustrate the concepts of conjugate elements and class. The implications of symmetry on transition dipole moments and transition polarizabilities lead to selection rules, which are explained and illustrated. The last section of this chapter consists of the analysis of symmetry properties of molecular vibrations in order to identify point groups. At the end of this chapter, the student can find several Mathematica codes (Classes of Symmetry Operations, Multiplication Table for the C_{2v} Group, Selection Rule for a Particle in a One-Dimensional Box, Selection Rule for the Two-Particle Rigid Rotor, Selection Rule for the Harmonic Oscillator) with references to the main text and detailed explanations of new commands, a glossary of important scientific terms, and a list of exercises, whose complete answers can be found after the Appendix.

4.1 Symmetry Operations

A symmetry operation converts the initial configuration of an object into an indistinguishable final configuration. If the symmetry operation is not observed during its application, it is impossible to decide whether it was applied, because the

J.J.C. Teixeira-Dias, *Molecular Physical Chemistry*,
DOI 10.1007/978-3-319-41093-7_4

initial and final configurations are indistinguishable. In this chapter, the objects acted on by symmetry operations are molecules with their nuclei at equilibrium positions. Each symmetry operation is carried out by a **symmetry element**. For a finite object such as a molecule, a symmetry element can be a plane, an axis, a point, or a combination of an axis and a plane. Symmetry operations and the corresponding symmetry elements are represented by the same letters, with symmetry operations shown in *italics*.

The symmetry of the H_2O molecule results from the equivalence of its OH bonds. There are three operations that express this symmetry (Fig. 4.1). The first is the anticlockwise rotation by $2\pi/2$ radians (180º) around the C_2 **axis of rotation** (this axis coincides with the z axis; anticlockwise is the direction of an angle in the plane). The second is the reflection through the xz **plane of reflection**, σ_{xz} (σ = sigma is the Greek letter for *s*; *Spiegel* is the German word for *mirror*). The third symmetry operation is the reflection in the yz plane of reflection σ_{yz}.

Consider now the benzene molecule at equilibrium, with two regular concentric and parallel hexagons, one formed by the carbon nuclei, the other formed by the hydrogen nuclei (Fig. 4.2). The most evident symmetry element of the benzene molecule is the sixfold rotation axis C_6, perpendicular to the reflection plane σ_h defined by the molecule's nuclei. The axis C_6 coincides with the rotation axes C_3 and C_2. As the highest-order axis, C_6 is the principal axis of the molecule. Reflection planes that contain C_6 are of two types, namely vertical reflection planes σ_v, which contain diametrically opposed CH bonds, and diagonal reflection planes σ_d, which bisect the angles formed by σ_v planes. If a Cartesian coordinate system is set up with its origin at the center of the regular hexagon formed by the carbon nuclei, the exchange of Cartesian coordinates (x, y, z) by $(-x, -y, -z)$ for any point in space leads to an indistinguishable molecular configuration (Fig. 4.2). This symmetry operation is called **inversion** and is represented by i, with the corresponding symmetry element being a point at the origin of the Cartesian reference system, called the **center of symmetry** i. Note that every inversion symmetry operation can be seen as resulting from a $2\pi/2$ rotation around one C_2 axis, followed by a reflection in σ_h. Hence, the center of symmetry i is said to be equivalent to a twofold **rotation–reflection axis** S_2. A symmetry operation around C_6 followed by

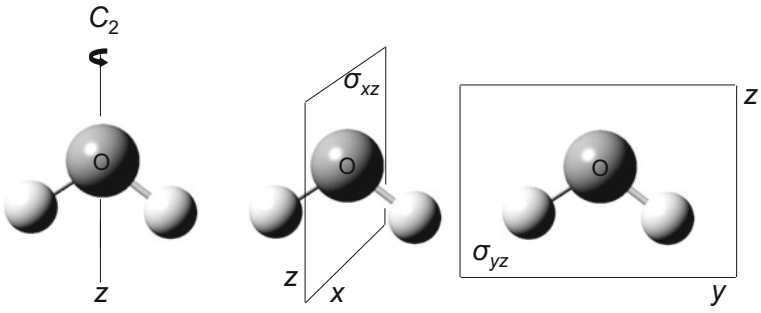


Fig. 4.1 Nontrivial symmetry elements of the H_2O molecule

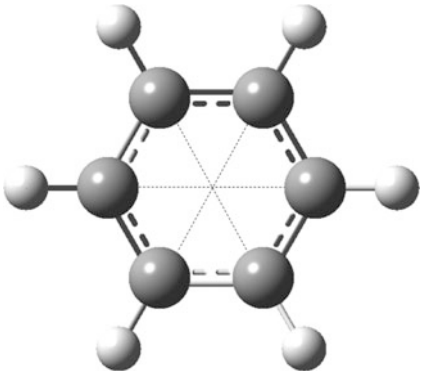


Fig. 4.2 The center of symmetry of a benzene molecule is at the center of the regular hexagon formed by the nuclei of the carbon atoms

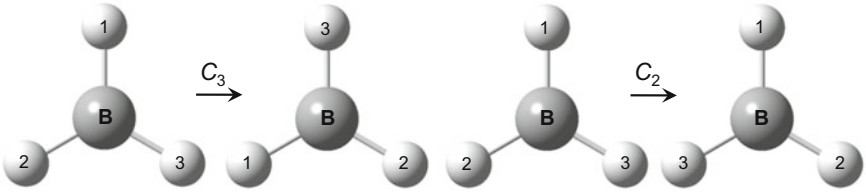


Fig. 4.3 C_3 and C_2 rotations in BF_3. Fluorine atoms are numbered in order to show the effect of the symmetry operation

a reflection in σ_h gives a sixfold rotation–reflection around the axis of symmetry S_6. In turn, combining C_3 with σ_h yields a rotation–reflection axis of symmetry S_3.

Let us now consider the BF_3 molecule. The structural equivalence of the B–F bonds, all in the same horizontal plane of reflection σ_h, points to the existence of a principal threefold rotation axis C_3 perpendicular to the molecular plane. In addition, each B–F bond axis is a twofold symmetry axis C_2 (Fig. 4.3). In turn, each perpendicular plane that contains one B–F bond axis is a vertical plane of symmetry, σ_v. A $2\pi/3$ rotation around the C_3 axis followed by a reflection in σ_h is a threefold rotation–reflection symmetry operation around S_3.

One of the symmetry elements of the CH_4 molecule is a fourfold rotation–reflection symmetry axis S_4 (Fig. 4.4). However, in this molecule, both the fourfold rotation axis C_4 and the perpendicular plane of reflection σ are not individual symmetry elements of CH_4.

A symmetry operation around a rotation–reflection axis S_1 involves a 2π rotation followed by a reflection in the perpendicular plane. These operations make S_1 equivalent to a reflection in a horizontal plane, $S_1 = \sigma_h$. In turn, an S_2 rotation–reflection axis involves a π rotation around the axis followed by a reflection in the perpendicular plane, the whole operation being equivalent to the inversion, $S_2 = i$, as we have mentioned before.

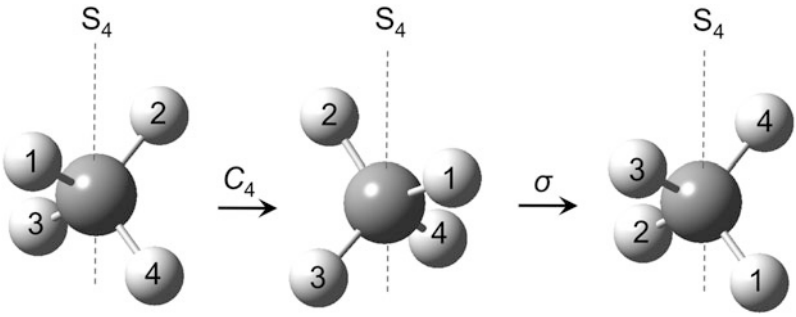


Fig. 4.4 The fourfold rotation–reflection symmetry operation on the CH_4 molecule is a $2\pi/4$ rotation (90º) followed by reflection in the perpendicular plane that contains the carbon nucleus

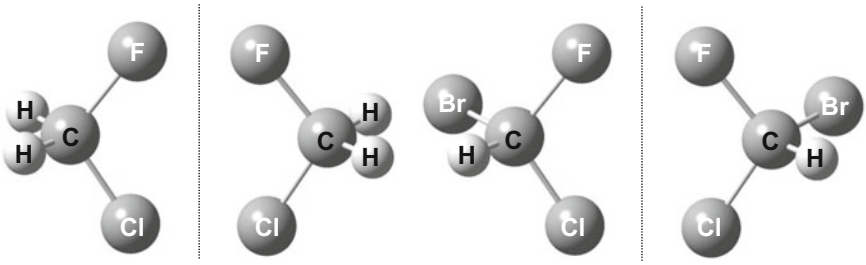


Fig. 4.5 Object-to-mirror image relationship for CH_2ClF and CHBrClF

Consider now the chlorofluoromethane (CH_2ClF) and bromochlorofluoromethane (CHBrClF) molecules. Figure 4.5 shows pairs of these molecules in the object–image relationship given by a plane mirror. The distinction between them is clear: while the two CH_2ClF molecules correspond to the *same* molecule, since a π rotation of one of them followed by a translation makes the object and image superimposable, the CHBrClF molecules are not superimposable. In fact, in order to convert one CHBrClF molecule into the other molecule of the pair, two bonds need to be broken and exchanged. The CHBrClF compound is said to be **chiral** (the word *chiral* has its origin in the Greek word for *hand*), since the molecules of the object–image pair are not superimposable, like the left and right hands of a person. From the symmetry point of view, chiral molecules do not have any symmetry rotation–reflection axis S_n.

4.2 Point Groups

If *A* and *B* represent two symmetry operations for the same molecule, the sequential application of these symmetry operations, *A* followed by *B*, defines the **product** of symmetry operations, denoted by *BA*, with *A* being applied first. Considering the definition of symmetry operation, we conclude that the product of two symmetry

operations is also a symmetry operation for the same molecule (the product is said to be an **internal** mathematical operation). In addition, the action corresponding to a particular symmetry operation can always be reversed. This means that every symmetry operation of a particular molecule has an **inverse**. Note that the product of a symmetry operation by its inverse is equivalent to "doing nothing," that is, to the **identity** E (from the German word *Einheit* for *unity*).

The set of symmetry operations for a particular molecule obeys the following requirements:

(1) the product of any two elements, including the product of an element and itself, is a member of the same set of symmetry operations (this is called the **closure** requirement);
(2) each symmetry operation has an **inverse**;
(3) there is an **identity** operation, which we represent by E;
(4) the product is **associative**, that is, $(AB)C = A(BC)$ (this property results from the concept of symmetry operation).

Having these properties, one says that the set of symmetry operations of a particular molecule belongs to a **point group** (all symmetry elements have a point in common which is kept unaltered by any symmetry operation of the group). The number of elements of a group is called the **order** of the group. When all products of symmetry operations are commutative, $BA = AB$, the group is said to be **commutative** or **abelian**. Group theory is a branch of mathematics.

We now consider a systematic way of finding the point group of a molecule. We begin by identifying groups of molecules with high symmetry, including linear molecules. These are generally called *special groups* (Fig. 4.6). The internuclear axis of linear molecules is a **principal axis** of infinite order C_∞, since no matter how small the rotation around the principal axis is, a new indistinguishable molecular configuration is always obtained. There are two groups for linear molecules, depending on whether they have a center of symmetry (Fig. 4.7). Linear molecules with no center of symmetry like HCl, HCN, and HC≡CCl (Fig. 4.7), belong to the $C_{\infty v}$ group, whereas linear molecules with a center of symmetry, like CO_2, HC≡CH (Fig. 4.7), and H_2, belong to the $D_{\infty h}$ group. Molecules with the shape of regular polyhedrons exhibit high symmetry and belong to the groups we classify as special groups. Regular tetrahedral molecules like CH_4, CCl_4 (Fig. 4.7), and P_4 belong to the T_d group (tetrahedral group). A regular octahedral molecule like SF_6 (Fig. 4.7) belongs to the O_h group (octahedral group).

If a molecule whose group of symmetry is to be identified does not belong to any of these groups of high symmetry, one should look for the presence of a C_n symmetry axis. A molecule that does not have a C_n axis of symmetry may belong to one of the three following groups: C_s, if the molecule, like HOCl (Fig. 4.8), has only one plane of symmetry; C_i, if the molecule, like CHClF-CHClF in the *trans* conformation (Fig. 4.8), has only one center of symmetry; C_1, if the molecule, like CHBrClF (Fig. 4.8), does not have any element of symmetry.

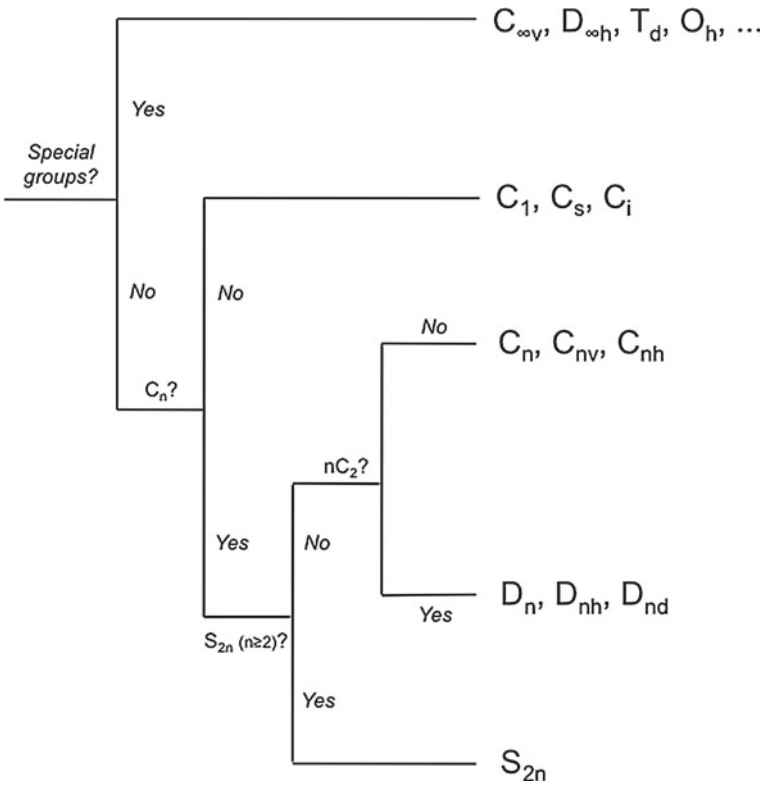


Fig. 4.6 Dichotomy for finding the point group of a molecule

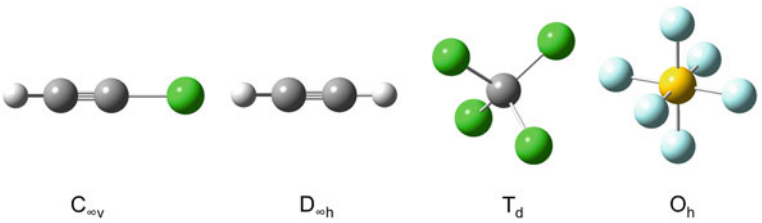


Fig. 4.7 Molecules with high symmetry

If the molecule does not belong to any special group or to any of the three last groups, it should have at least one C_n symmetry axis (possibly n = 2). Then, one should ask whether an S_{2n} symmetry element ($n \geq 2$) collinear with the C_n axis of symmetry is present. If the answer to this question is no, then one should ask whether there are nC_2 axes of symmetry perpendicular to the C_n axis of symmetry. Answering no to this question leads us to three groups of symmetry, namely, C_n (HOOH as an example of the C_2 group; Fig. 4.9), C_{nh} (*trans*-FHC = CHF as an

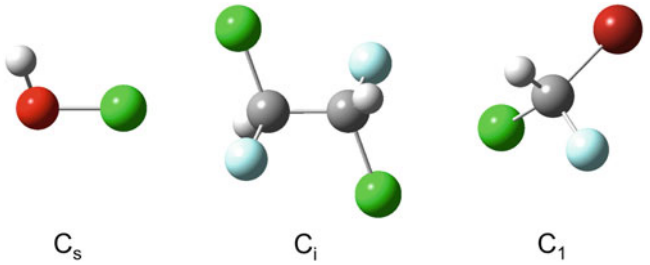


Fig. 4.8 Molecules that do not have a C_n axis of symmetry

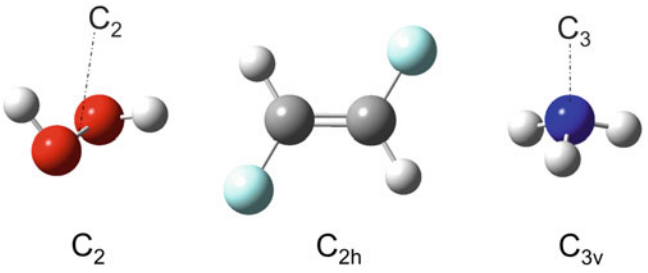


Fig. 4.9 Molecules with a principal axis of symmetry C_n but no C_2 perpendicular axes

example of the C_{2h} group; Fig. 4.9); and C_{nv} (NH_3 as an example of the C_{3v} group; Fig. 4.9).

Molecules with nC_2 axes of symmetry perpendicular to the C_n axis of symmetry belong to groups D_n [tris(ethylenediamine)cobalt (III) cation, $[Co(en)_3]^{3+}$ with en = $H_2NCH_2CH_2NH_2$, as an example of the D_3 group; Fig. 4.10], D_{nh} (ethylene as an example of the D_{2h} group; Fig. 4.10) and D_{nd} (allene as an example of the D_{2d} group; Fig. 4.10).

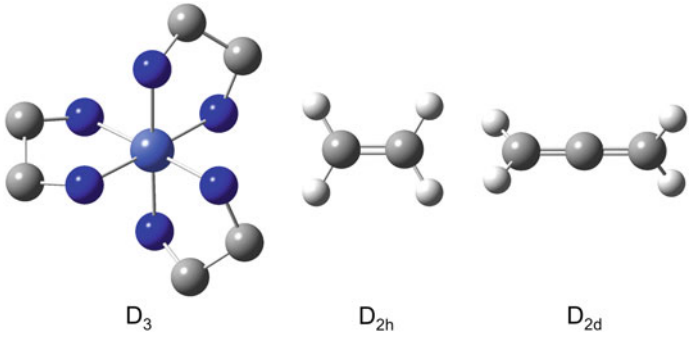


Fig. 4.10 Molecules with nC_2 axes of symmetry perpendicular to the principal axis of symmetry

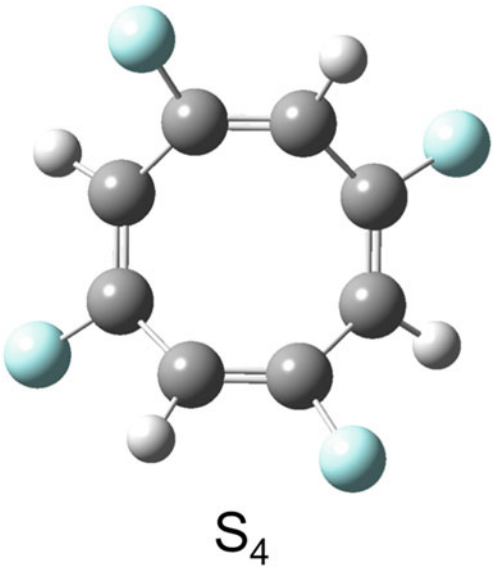


Fig. 4.11 The 1,3,5,7-tetrafluorooctatetraene molecule

If the answer to the question in the presence of an S_{2n} symmetry element ($n \geq 2$) is yes, then the molecule belongs to an S_{2n} group (the 1,3,5,7-tetrafluorooctatetraene molecule is shown in Fig. 4.11 as an example of the S_4 group) (**E1**).

4.3 Matrix Representations

A transformation that converts the vector $(x, y, z)^T$ to $(x', y', z')^T$ (the superscript T stands for "transpose," thus giving the corresponding column vector) is mathematically represented by a matrix whose elements are the coefficients of the linear combinations of the original coordinates. Such a transformation is linear and applies within the same vector space. Let us find the representation matrices for symmetry operations, namely the identity, reflection, inversion, rotation, and rotation–reflection. Starting with the identity E and the column vector $(x, y, z)^T$, we can write

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = E \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{4.1}$$

since this operation "does nothing," that is, $x' = x$, $y' = y$, and $z' = z$. One says that the 3×3 identity matrix represents the identity operation and the column vector $(x, y, z)^T$ is the **basis** for this representation.

For the reflection through the xz plane, σ_{xz}, we write

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \sigma_{xz} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{4.2}$$

because $x' = x$, $y' = -y$, and $z' = z$. Inversion is given by the equality

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = i \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{4.3}$$

since $x' = -x$, $y' = -y$, and $z' = -z$.

Figure 4.12 represents the counterclockwise rotation of a vector in the xy plane by an angle θ around the z-axis starting from an angle α with the positive x-axis. Simple trigonometry enables us to write

$$x = r \cos \alpha \qquad y = r \sin \alpha \tag{4.4}$$

and

$$\begin{aligned} x' &= r \cos(\alpha + \theta) = r(\cos \alpha \cos \theta - \sin \alpha \sin \theta) = x \cos \theta - y \sin \theta \\ y' &= r \sin(\alpha + \theta) = r(\cos \alpha \sin \theta + \sin \alpha \cos \theta) = x \sin \theta + y \cos \theta \end{aligned} \tag{4.5}$$

Hence, in matrix notation, we can write

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{4.6}$$

Considering now the basis formed by the column vector $(x, y, z)^T$, we write

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = C_n \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{4.7}$$

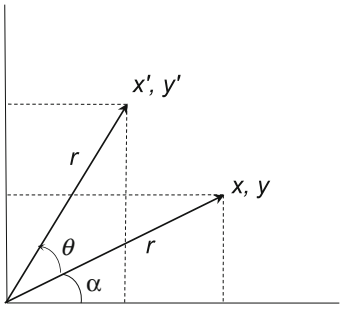


Fig. 4.12 Counterclockwise rotation around the z axis of a vector by an angle θ in the xy plane

where C_n represents the counterclockwise rotation of a vector in the xy plane by an angle θ around the z-axis that coincides with the C_n axis (**E2**).

The rotation–reflection operation S_n is given by the product $\sigma_{xy}C_n$ (the plane of reflection is perpendicular to the C_n axis, which is coincident with the z-axis), and so we can write

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = S_n \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
$$= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{4.8}$$

(**E3**, **E4**, **E5**).

We now consider the C_{2v} point group. The matrix that represents C_2 can be easily obtained from (4.7) using $\theta = 2\pi/2$ radians, and the matrices that represent the reflections are diagonal, because the planes of reflection coincide with xz and yz planes. Hence,

$$\underset{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}{E} \quad \underset{\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}{C_2} \quad \underset{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}{\sigma_{xz}} \quad \underset{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}{\sigma_{yz}} \tag{4.9}$$

We now pass to the C_{3v} point group and consider the group of nontrivial symmetry operations C_3, C_3^2 (= C_3C_3), σ_{v1}, σ_{v2}, and σ_{v3}. Assuming that C_3 coincides with the z-axis and σ_{v1} uses yz as the plane of reflection (Fig. 4.13), then C_3 and C_3^2 are represented by the following 3×3 matrices [$\theta(C_3) = 2\pi/3$ radians and θ (C_3^2) = $4\pi/3$ radians, see (4.7)],

$$\underset{\begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}}{C_3} \quad \underset{\begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}}{C_3^2} \tag{4.10}$$

and σ_{v1} is represented by

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.11}$$

We now consider a new basis $(x'\ y'\ z')^T$ obtained from $(x\ y\ z)^T$ by a counterclockwise rotation of $2\pi/3$ radians around the z-axis that is coincident with both z' and C_3 (Fig. 4.13):

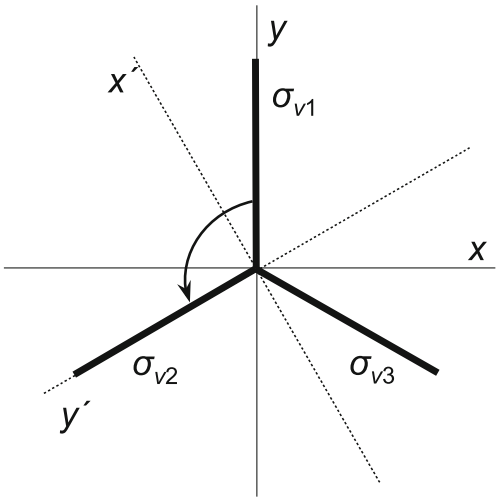


Fig. 4.13 Horizontal projections of vertical planes of the C_{3v} point group

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = C_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{4.12}$$

Inspection of Fig. 4.13 shows that the change of basis (4.12) converts σ_{v1} into σ_{v2}. Therefore, σ_{v2} is obtained from σ_{v1} by a similarity transformation $B = X^{-1}A\,X$, where X is C_3 [σ_{v1} and σ_{v2} are represented by similar matrices; see (2.89)]:

$$\sigma_{v2} = C_3^{-1}\sigma_{v1}C_3 \tag{4.13}$$

Following the same reasoning, σ_{v3} is obtained from σ_{v1} by another similarity transformation in which X is C_3^2 (the change of basis is due to C_3^2):

$$\sigma_{v3} = \left(C_3^2\right)^{-1}\sigma_{v1}C_3^2 \tag{4.14}$$

Note that $C_3^{-1} = C_3^T$ and $\left(C_3^2\right)^{-1} = \left(C_3^2\right)^T$, since both $(x\ y\ z)^T$ and $(x'\ y'\ z')^T$ are orthogonal systems of axes [see (4.7)]. Substitution of (4.10) and (4.11) in (4.13) and (4.14) leads to

$$\overset{\sigma_{v2}}{\begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}} \quad \overset{\sigma_{v3}}{\begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}} \tag{4.15}$$

Hence, the matrices that represent the symmetry operations of the C_{3v} group in the basis $(x\ y\ z)^T$ are [see (4.10), (4.11), and (4.15)]

$$\begin{matrix} E \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad \begin{matrix} C_3 \\ \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad \begin{matrix} C_3^2 \\ \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad \begin{matrix} \sigma_{v1} \\ \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad \begin{matrix} \sigma_{v2} \\ \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad \begin{matrix} \sigma_{v3} \\ \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix} \tag{4.16}$$

In addition, C_3 and C_3^2 are also related by similarity transformations $B = X^{-1}A\,X$, where X can be σ_{v1}, σ_{v2}, or σ_{v3}. A similarity transformation preserves the sum of the diagonal matrix elements, or trace; in group theory, **character** (**E6**). The character of the identity symmetry operation E is always equal to the dimension of the basis. The characters of the symmetry operations of the C_{3v} point group in the basis $(x\ y\ z)^T$ are given by

$$\begin{matrix} E & C_3 & C_3^2 & \sigma_{v1} & \sigma_{v2} & \sigma_{v3} \\ 3 & 0 & 0 & 1 & 1 & 1 \end{matrix} \tag{4.17}$$

The symmetry operations of the C_{3v} point group (E, C_3, C_3^2, σ_{v1}, σ_{v2}, σ_{v3}) can be grouped in three sets, $\{E\}$, $\{C_3, C_3^2\}$, and $\{\sigma_{v1}, \sigma_{v2}, \sigma_{v3}\}$, formed by symmetry operations having the same characters and related by similarity transformations. Each of these sets of symmetry operations is called a **class**, and the symmetry operations within one specified class are said to be mutually **conjugate** (**E7**). Character tables group symmetry operations in classes. Using matrix representations of the symmetry operations for the C_{3v} point group in the basis $(x\ y\ z)^T$ [see (4.10), (4.11), (4.15)], the first *Mathematica* code of **M1** confirms that $\{C_3, C_3^2\}$, and $\{\sigma_{v1}, \sigma_{v2}, \sigma_{v3}\}$ are two classes of the C_{3v} point group, and the second code of **M1** enables one to observe $\sigma_{v1}^T.C_3(\theta).\sigma_{v1}$ "behind the scenes."

4.4 Character Tables

The product of two symmetry operations of a point group is a member of the same point group. The multiplication table of a point group presents the results of the products of symmetry operations BA, where A and B are symmetry operations of the group and A is applied first (A and B are presented as columns and rows of the multiplication table). Using matrix representations of the symmetry operations of the C_{2v} point group in the basis $(x\ y\ z)^T$ [see (4.9)], the *Mathematica* code **M2** builds the multiplication table of the C_{2v} point group. Inspection of this multiplication table shows that each symmetry operation of the C_{2v} group coincides with its own inverse, since the diagonal is occupied by the identity. In addition, the upper and lower matrix triangles of the multiplication table are equal (the product matrix is symmetric), meaning that the order of multiplication in a product is irrelevant; that is, symmetry operations commute, and C_{2v} is an abelian point group. Among the nontrivial symmetry operations C_2, σ_{xz}, and σ_{yz}, the product of two of them is equal

to the one left out. A set of matrices or numbers (a number is a 1×1 matrix) consistent with the multiplication table is called a **group representation**.

Multiplication table of the C_{2v} group

	E	C_2	σ_{xz}	σ_{yz}
E	E	C_2	σ_{xz}	σ_{yz}
C_2	C_2	E	σ_{yz}	σ_{xz}
σ_{xz}	σ_{xz}	σ_{yz}	E	C_2
σ_{yz}	σ_{yz}	σ_{xz}	C_2	E

Consider now a molecular orbital for H_2O. The multiplication table for the C_{2v} symmetry group can be used to find how a particular molecular orbital transforms when a symmetry operation is applied to it. Applying the identity operator to a function does not modify it in any conceivable way, the result being the same as if the function had been multiplied by 1, $E\,\phi_i = \phi_i$. Consider now the diagonal of the C_{2v} multiplication table: the identity operation can be replaced by the square of any other symmetry operation, that is,

$$E = C_2C_2 = \sigma_{xz}\sigma_{xz} = \sigma_{yz}\sigma_{yz} \tag{4.18}$$

Applying each member of (4.18) to a function, one obtains

$$E\phi_i = C_2C_2\phi_i = \sigma_{xz}\sigma_{xz}\phi_i = \sigma_{yz}\sigma_{yz}\phi_i = \phi_i \tag{4.19}$$

Therefore, the eigenvalues for the nontrivial symmetry operations of H_2O (C_2, σ_{xz}, σ_{yz}) should be either +1 or −1. Since the product of two of the nontrivial symmetry operations (C_2, σ_{xz}, σ_{yz}) is equal to the one left out [this can be confirmed by inspection of the elements (2,3) (second row, third column), (2,4), (3,2), (3,4), (4,2), and (4,3) of the multiplication table], either all of the corresponding eigenvalues of the C_{2v} point group have to be equal to +1, or two of them are equal to –1. Each set of these eigenvalues consistent with the multiplication table of the group is called an **irreducible representation** of the C_{2v} group (the word "irreducible" means that these representations cannot be further decomposed into simpler ones). The table that contains the characters of the irreducible representations of the group is called the **character table**.

The irreducible representations with 1 in the axis of highest order (the so-called **principal axis**; for the C_{2v} group, the C_2 axis) are identified by A. In turn, those representations with −1 in the principal axis are identified by B. The subscript 1 is used when the eigenvalue of the symmetry operation σ_{xz} is equal to +1. The representation with 1 as eigenvalues in all symmetry operations of the group reflects the maximum symmetry within the group and is called a **totally symmetric species**.

Character table of the C_{2v} point group

	E	C_2	σ_{xz}	σ_{yz}		
A_1	1	1	1	1	z	x^2, y^2, z^2
A_2	1	1	−1	−1	R_z	xy
B_1	1	−1	1	−1	x, R_y	xz
B_2	1	−1	−1	1	y, R_x	yz

(http://symmetry.jacobs-university.de)

The character table also presents the functions x, y, z, x^2, y^2, z^2, xy, xz, yz, and the rotations R_x, R_y, and R_z around the x-, y-, and z-axes. The z-coordinate transforms like the totally symmetric species (A_1), since the z-axis coincides with C_2 (Fig. 4.14). Inspection of Fig. 4.14 for the orientation of the x- and y-axes shows that x transforms like B_1 and y transforms like B_2. The x, y and z functions represent translations along these axes and so transform like the dipole moment components. The function xy transforms like A_2 ($= B_1 \times B_2$), xz and yz transform like B_1 ($= B_1 \times A_1$) and B_2 ($= B_2 \times A_1$), respectively, and x^2 ($B_1 \times B_1 = A_1$), y^2 ($B_2 \times B_2 = A_1$), and z^2 ($A_1 \times A_1 = A_1$) are totally symmetric. Under symmetry operations, the polarizability matrix elements transform according to the corresponding subscripts (α_{xx} transforms as x^2, α_{xy} transforms as xy, …).

The rotations around the x-, y-, and z-axes are carried out by the corresponding torques (Fig. 4.15). In particular, by applying the symmetry operations of C_{2v} to the torque for R_x, we conclude that R_x transforms like the symmetry species B_2 (see Fig. 4.15 and the character table). Similar reasoning applied to the torques for R_y and R_z leads to their assignments to B_1 and A_2, respectively.

Inspection of the C_{2v} character table reveals that the irreducible representations behave like "orthogonal vectors." In fact, the dot product of two irreducible representations is given by

$$\Gamma_i^T \cdot \Gamma_j = h\,\delta_{ij} \tag{4.20}$$

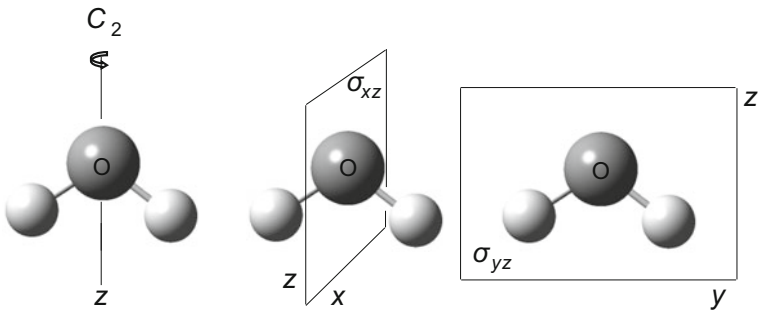


Fig. 4.14 Nontrivial symmetry elements of the H_2O molecule

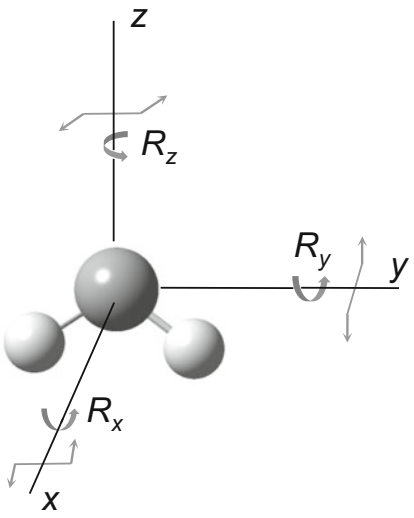


Fig. 4.15 Rotations R_x, R_y, and R_z for the H_2O molecule (C_{2v} group) and the corresponding torques

where the character of each irreducible representation is represented by a column vector, h is the order of the group (= number of symmetry elements; for C_{2v}, $h = 4$), and δ_{ij} is the Kronecker delta ($\delta_{ij} = 0$ for $i \neq j$, $\delta_{ij} = 1$ for $i = j$).

The character of matrix representation (4.9) (= set of characters of the symmetry operations) is given by

$$\begin{array}{cccccc} C_{2v} & E & C_2 & \sigma_{xz} & \sigma_{yz} \\ \Gamma & 3 & -1 & 1 & 1 \end{array} \tag{4.21}$$

A reducible representation Γ can be decomposed as a linear combination of irreducible representations of the group

$$\Gamma = \sum a_i \Gamma_i \tag{4.22}$$

where

$$a_i = \frac{\Gamma^T \cdot \Gamma_i}{h} \tag{4.23}$$

since substitution of (4.22) into (4.23) using (4.20) gives an identity. Applying (4.23) to Γ in (4.21) leads to

$$\Gamma = \mathrm{A}_1 + \mathrm{B}_1 + \mathrm{B}_2 \tag{4.24}$$

a result that is not at all surprising, since A_1, B_1, and B_2 are the irreducible representations for the functions x, y, and z that form the basis for Γ (see the character table of the C_{2v} point group).

We now consider the character table for the C_{3v} group (http://symmetry.jacobs-university.de) and note that the irreducible representation of the point group labeled E is two-dimensional, since it is shown by the character of the identity operation E and the basis functions grouped in pairs, on the right-hand side of the irreducible representation E (**E8**).

The characters of the symmetry operations of the C_{3v} point group in the $(x\ y\ z)^T$ basis [see (4.16)] are given by

$$\begin{array}{cccc} C_{3v} & E & 2C_3 & 3\sigma_v \\ \Gamma & 3 & 0 & 1 \end{array} \tag{4.25}$$

Use of (4.20), (4.22), and (4.23) leads to

$$\Gamma = A_1 + E \tag{4.26}$$

a result that can be easily confirmed by inspection of the character table for the C_{3v} point group.

Character table of the C_{3v} point group

	E	$2C_3$	$3\sigma_v$		
A_1	1	1	1	z	$x^2 + y^2, z^2$
A_2	1	1	−1	R_z	
E	2	−1	0	(x, y) (R_x, R_y)	$(x^2–y^2, xy)$ (xz, yz)

(http://symmetry.jacobs-university.de)

4.5 Selection Rules

A function $f(x)$ is said to be **even** if $f(x) = f(-x)$. Functions like x^2, cos x, $\exp(-x^2)$, $|x|$ are even functions. In turn, $g(x)$ is said to be an **odd function** if $g(x) = -g(-x)$. Functions like x, x^3, sin x are odd functions. Let us now consider the integral of a function $h(x)$ defined for $-a \leq x \leq a$. It is easy to show that $h(x)$ can be decomposed into the sum of the even function $f(x) = [\ h(x) + h(-x)]/2$ and the odd function $g(x) = [\ h(x) - h(-x)]/2$. If the odd function is integrable, then its integral over the symmetric interval $[-a, a]$ is identically zero. Thus, we can conclude that only the even function contributes to the integral of $f(x)$ over the symmetric interval $[-a, a]$, with this integral being precisely the same as twice the integral over the interval $[0, a]$. Generalization of this result enables one to conclude that the integral of $f(s)$ over the entire configuration space is significant when $f(s)$ is a basis for a

representation that coincides with or includes the totally symmetric irreducible representation of the group.

We now recall a result obtained in the previous chapter [see (3.283)] showing that the probability of a spectroscopic transition in absorption from stationary state i to stationary state f is proportional to

$$\left|\left\langle \psi_f^0|\hat{\mu}_x|\psi_i^0\right\rangle\right|^2+\left|\left\langle \psi_f^0|\hat{\mu}_y|\psi_i^0\right\rangle\right|^2+\left|\left\langle \psi_f^0|\hat{\mu}_z|\psi_i^0\right\rangle\right|^2=\left|\left\langle \psi_f^0|\hat{\boldsymbol{\mu}}|\psi_i^0\right\rangle\right|^2 \tag{4.27}$$

The spectroscopic transition from i to f is **active** if at least one of the transition dipole moment integrals is different from zero. Assuming that this applies to the first term of the sum (4.27), then the product of functions

$$\left(\psi_f^0\right)^* x\,\psi_i^0 \tag{4.28}$$

belongs to a representation that coincides with or contains the totally symmetric representation of the group. In other words, the product of functions

$$\left(\psi_f^0\right)^* \psi_i^0 \tag{4.29}$$

is the basis for a representation that coincides with or contains the irreducible representation for the function x.

When

$$\left\langle \psi_f^0|\hat{\boldsymbol{\mu}}|\psi_i^0\right\rangle = 0 \tag{4.30}$$

the transition from i to f is **forbidden**. **Allowed** transitions have

$$\left\langle \psi_f^0|\hat{\boldsymbol{\mu}}|\psi_i^0\right\rangle \neq 0 \tag{4.31}$$

Consider a charged particle in a one-dimensional box (see 3.**§1**). When electromagnetic radiation is absorbed or emitted, the corresponding **selection rule** indicates that the quantum number must change by an odd integer, as illustrated by the *Mathematica* code **M3**. In turn, the *Mathematica* code **M4** verifies that the selection rule for a two-particle rigid rotor (see 3.**§2**), with an associated dipole moment for the electromagnetic radiation to interact with, is $\Delta J = \pm 1$, where J is the rotational quantum number. In addition, the selection rule for a harmonic oscillator (see Sect. 3.2) indicates that for absorption and stimulated emission of radiation, one has $\Delta n = \pm 1$, where n is the vibrational quantum number (see *Mathematica* code **M5**).

Considering now a Raman scattering transition from i to f, we recall that the xy polarizability component is given by

$$(\alpha_{xy})_{fi} = \frac{1}{\hbar}\sum_{r\neq i,f}\left(\frac{\left\langle\psi_f^0|\hat{\boldsymbol{\mu}}_x|\psi_r^0\right\rangle\left\langle\psi_r^0\middle|\hat{\boldsymbol{\mu}}_y\middle|\psi_i^0\right\rangle}{\omega_{ri}-\omega_0-i\Gamma_r}+\frac{\left\langle\psi_f^0\middle|\hat{\boldsymbol{\mu}}_y\middle|\psi_r^0\right\rangle\langle\psi_r^0|\hat{\boldsymbol{\mu}}_x|\psi_i^0\rangle}{\omega_{rf}+\omega_0+i\Gamma_r}\right) \tag{4.32}$$

This expression suggests that the Raman selection rule depends on the product of transition dipole moments in the numerators of (4.32),

$$\left\langle\psi_f^0|x|\psi_r^0\right\rangle\langle\psi_r^0|y|\psi_i^0\rangle \tag{4.33}$$

and involve a summation over all virtual states r. However, since no direct information can be obtained about the virtual states r, except perhaps in discrete resonance Raman, where a limited number of them dominate the sum over r, the general condition for

$$(\alpha_{xy})_{fi} = \left\langle\psi_f^0|\alpha_{xy}|\psi_i^0\right\rangle \tag{4.34}$$

to be nonzero is that

$$\left\langle\psi_f^0|xy|\psi_i^0\right\rangle \tag{4.35}$$

belongs to a representation that contains the totally symmetric species, a condition that can be easily extended to other components of the polarizability tensor.

4.6 Molecular Vibrations

We now determine the symmetry species of the vibrations of a molecule and find the numbers of distinct fundamental frequencies in infrared and Raman spectra. In small molecules, these numbers may eventually be used to distinguish isomers or identify geometries corresponding to a known molecular stoichiometry.

We begin by considering the *cis*-1,2-dichloroethene molecule (Fig. 4.16). This molecule belongs to the C_{2v} point group, whose character table is

Character table of the C_{2v} point group

	E	C_2	σ_{xz}	σ_{yz}		
A_1	1	1	1	1	z	x^2, y^2, z^2
A_2	1	1	−1	−1	R_z	xy
B_1	1	−1	1	−1	x, R_y	xz
B_2	1	−1	−1	1	y, R_x	yz

(http://symmetry.jacobs-university.de)

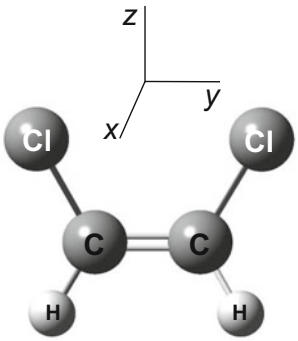


Fig. 4.16 *cis*-1,2-Dichloroethene (C_{2v} group) with system of Cartesian axes. The latter should be centered on each nucleus of the molecule, thus yielding a basis of 18 coordinates

The set of unit vectors directed along the Cartesian axes and centered on each nucleus of the molecule forms a basis of 18 vectors (= 3 N, where $N = 6$ is the number of atoms in the molecule) that can describe all molecular motions, namely translations, rotations, and vibrations. Therefore, these vectors form the basis for a reducible representation of the C_{2v} group. The diagonal elements of the matrices that represent the symmetry operations can be easily obtained without the need for determining the whole representation matrices. Simple and intuitive rules can be applied. In particular, when a vector moves its origin under the action of a symmetry operation, one should count zero for the diagonal element of the corresponding transformation matrix. In turn, if a vector remains unaltered under the action of a symmetry operation, one should count 1. When a vector rotates by θ radians, the trace of the corresponding matrix is given by $2\cos\theta + 1$ [see (4.7)]. Thus, applying the symmetry operations to the set of the basis vectors (see Fig. 4.16) gives the following reducible representation:

$$\begin{array}{cccccc} C_{2v} & E & C_2 & \sigma_{xz} & \sigma_{yz} \\ \Gamma & 18 & 0 & 0 & 6 \end{array} \tag{4.36}$$

Using (4.22) and (4.23), we decompose Γ into the irreducible representations of the group, obtaining

$$\Gamma = 6A_1 + 3A_2 + 3B_1 + 6B_2 \tag{4.37}$$

Inspection of the character table of the C_{2v} group shows that the irreducible representations of the group that correspond to rotations R_x, R_y, and R_z are B_2, B_1, and A_2, and those that correspond to translations x, y, and z are B_1, B_2, and A_1. Thus, we can write

$$\begin{aligned}\Gamma_{\text{rot}} &= A_2 + B_1 + B_2 \\ \Gamma_{\text{trans}} &= A_1 + B_1 + B_2\end{aligned} \tag{4.38}$$

Subtracting (4.38) from (4.37) leads to the reducible representation for the molecular vibrations

$$\Gamma_{\text{vib}} = 5A_1 + 2A_2 + B_1 + 4B_2 \tag{4.39}$$

Irreducible representations that transform like x, y, and z (A_1, B_1, B_2) correspond to infrared active vibrations, and those that transform like x^2, y^2, z^2, xy, xz, and yz (A_1, A_2, B_1, B_2) correspond to Raman active vibrations. Hence,

$$\begin{aligned}\Gamma_{\text{IR}} &= 5A_1 + B_1 + 4B_2 \\ \Gamma_{\text{R}} &= 5A_1 + 2A_2 + B_1 + 4B_2\end{aligned} \tag{4.40}$$

Therefore, the *cis*-1,2-dichloroethene molecule has 10 infrared active fundamental vibrations and 12 Raman active fundamental vibrations.

We now consider the *trans*-1,2-dichloroethene molecule (Fig. 4.17) that belongs to the C_{2h} group, whose character table is

Character table of the C_{2h} group

	E	C_2	i	σ_h		
A_g	1	1	1	1	R_z	x^2, y^2, z^2, xy
B_g	1	−1	1	−1	R_x, R_y	xz, yz
A_u	1	1	−1	−1	z	
B_u	1	−1	−1	1	x, y	

(http://symmetry.jacobs-university.de)

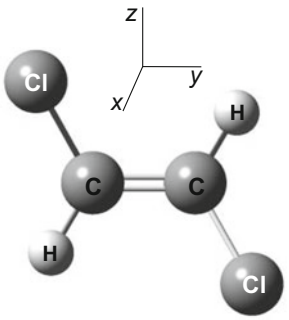


Fig. 4.17 *trans*-1,2-Dichloroethene (C_{2h} group) with system of Cartesian axes. The latter should be centered on each nucleus of the molecule thus yielding a basis of 18 coordinates

In the labeling of the irreducible representations, the subscripts g (from German *gerade*, *even*) and u (from German *ungerade*, *odd*) are used when the character for the inversion is positive and negative, respectively.

The set of unit vectors directed along the Cartesian axes and centered on each nucleus of the molecule forms a basis of 18 vectors and gives the following reducible representation:

$$\begin{array}{lccccc} \mathrm{C_{2h}} & E & C_2 & i & \sigma_h \\ \Gamma & 18 & 0 & 0 & 6 \end{array} \tag{4.41}$$

Using (4.22) and (4.23), we decompose Γ into the irreducible representations of the group, obtaining

$$\Gamma = 6\mathrm{A}_g + 3\mathrm{B}_g + 3\mathrm{A}_u + 6\mathrm{B}_u \tag{4.42}$$

Inspection of the character table reveals that

$$\begin{aligned} \Gamma_{\mathrm{rot}} &= \mathrm{A}_g + 2\mathrm{B}_g \\ \Gamma_{\mathrm{trans}} &= \mathrm{A}_u + 2\mathrm{B}_u \end{aligned} \tag{4.43}$$

We now subtract (4.43) from (4.42) to obtain the reducible representation for the molecular vibrations

$$\Gamma_{vib} = 5\mathrm{A}_g + \mathrm{B}_g + 2\mathrm{A}_u + 4\mathrm{B}_u \tag{4.44}$$

From the character table, we can conclude that the symmetry species of the infrared active and Raman active vibrations are

$$\begin{aligned} \Gamma_{\mathrm{IR}} &= 2\mathrm{A}_u + 4\mathrm{B}_u \\ \Gamma_{\mathrm{R}} &= 5\mathrm{A}_g + \mathrm{B}_g \end{aligned} \tag{4.45}$$

The symmetry species for the infrared active vibrations are of the u-type, whereas the symmetry species for the Raman active vibrations are of the g-type (**E9**).

Inspection of the character tables leads one to conclude that those groups that include inversion as a symmetry operation have **mutual exclusion** of the infrared and Raman active vibrations, i.e., the infrared active vibrations are Raman inactive; and conversely, the Raman active vibrations are infrared inactive. Comparison of (4.40) with (4.45) leads one to conclude that the *cis* and *trans* isomers can be distinguished by the number of infrared active and Raman active fundamental vibrations.

Consider now a molecule of general formula $\mathrm{XY_3}$, with the X atom bonded to each of the Y atoms. For this molecular connectivity, the molecule can be pyramidal ($\mathrm{C_{3v}}$ group) or planar ($\mathrm{D_{3h}}$ group). As we will find, this question can be answered by symmetry analysis of vibrations, in particular, by the number of distinct fundamental frequencies.

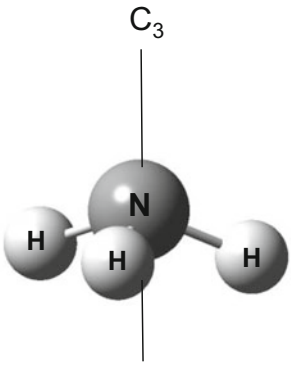


Fig. 4.18 A pyramidal XY_3 molecule includes, as symmetry elements, a threefold rotation axis C_3 and three vertical planes of reflection

A pyramidal XY_3 molecule (Fig. 4.18) belongs to the C_{3v} point group, whose character table is given by

Character table of the C_{3v} group

C_{3v}	E	$2C_3$	$3\sigma_v$		
A_1	1	1	1	z	$x^2 + y^2, z^2$
A_2	1	1	−1	R_z	
E	2	−1	0	(x, y) (R_x, R_y)	$(x^2–y^2, xy)$ (xz, yz)

(http://symmetry.jacobs-university.de)

We now consider the set of unit vectors directed along the Cartesian axes and centered on each nucleus of the XY_3 molecule. These form a basis of 12 vectors that correspond to the following reducible representation:

$$\begin{array}{cccc} C_{3v} & E & 2C_3 & 3\sigma_v \\ \Gamma & 12 & 0 & 2 \end{array} \tag{4.46}$$

Decomposing Γ into the irreducible representations of the group leads to

$$\Gamma = 3A_1 + A_2 + 4E \tag{4.47}$$

and inspection of the character table reveals that

$$\begin{array}{l} \Gamma_{rot} = A_2 + E \\ \Gamma_{trans} = A_1 + E \end{array} \tag{4.48}$$

Hence we have

$$\Gamma_{\text{vib}} = 2A_1 + 2E \tag{4.49}$$

and the following infrared active and Raman active vibrations:

$$\begin{aligned} \Gamma_{\text{IR}} &= 2A_1 + 2E \\ \Gamma_{\text{R}} &= 2A_1 + 2E \end{aligned} \tag{4.50}$$

Thus, a pyramidal XY_3 molecule (C_{3v} group) has four distinct fundamental frequencies of vibration, all infrared and Raman active. Note that each E species corresponds to two degenerate vibrational modes and a single distinct frequency.

The planar XY_3 molecule (Fig. 4.19) belongs to the D_{3h} group, whose character table is

D_{3h}	E	$2C_3$	$3C_2$	σ_h	$2S_3$	$3\sigma_v$		
A'_1	1	1	1	1	1	1		$x^2 + y^2, z^2$
A'_2	1	1	−1	1	1	−1	R_z	
E'	2	−1	0	2	−1	0	(x, y)	$(x^2 - y^2, xy)$
A''_1	1	1	1	−1	−1	−1		
A''_2	1	1	−1	−1	−1	1	z	
E''	2	−1	0	−2	1	0	(R_x, R_y)	(xz, yz)

(http://symmetry.jacobs-university.de)

The set of unit vectors directed along the Cartesian axes and centered on each nucleus of the XY_3 molecule form a basis of 12 vectors that correspond to the following reducible representation:

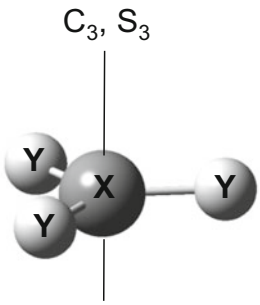


Fig. 4.19 A planar XY_3 molecule includes, as symmetry elements, a threefold rotation axis C_3 that coincides with a threefold rotation-reflection axis S_3, three twofold rotation axes C_2 that are aligned with each of the X–Y bonds, and four planes of symmetry: one horizontal plane and three vertical planes

$$\begin{array}{lccccccc} \mathrm{D}_{3h} & E & 2C_3 & 3C_2 & \sigma_h & 2S_3 & 3\sigma_v \\ \Gamma & 12 & 0 & -2 & 4 & -2 & 2 \end{array} \tag{4.51}$$

Expressing this reducible representation as a linear combination of irreducible representations of the group leads to

$$\Gamma = \mathrm{A}_1' + \mathrm{A}_2' + 3\mathrm{E}' + 2\mathrm{A}_2'' + \mathrm{E}'' \tag{4.52}$$

The symmetry species for rotations and translations are

$$\begin{array}{l} \Gamma_{\mathrm{rot}} = \mathrm{A}_2' + \mathrm{E}'' \\ \Gamma_{\mathrm{trans}} = \mathrm{E}' + \mathrm{A}_2'' \end{array} \tag{4.53}$$

Hence,

$$\Gamma_{\mathrm{vib}} = \mathrm{A}_1' + 2\mathrm{E}' + \mathrm{A}_2'' \tag{4.54}$$

and

$$\begin{array}{l} \Gamma_{\mathrm{IR}} = 2\mathrm{E}' + \mathrm{A}_2'' \\ \Gamma_{\mathrm{R}} = \mathrm{A}_1' + 2\mathrm{E}' \end{array} \tag{4.55}$$

In conclusion, a planar XY_3 molecule (D_{3h} group) has three distinct fundamental frequencies of vibration, both in the infrared and Raman spectra. Thus, it is possible to distinguish the geometries of an XY_3 molecule by the number of distinct fundamental frequencies of vibration, since this result differs from that obtained for a pyramidal XY_3 molecule [see (4.50)].

In Raman spectroscopy, the incident radiation is a plane-polarized electromagnetic wave. Analyzing the polarization of the scattered light may provide useful information on the symmetry of vibrations. In order to carry on this polarization analysis, we need to interpose a polarization analyzer in the way of the scattered light, just before this enters the spectrometer. This polarization analyzer acts like a filter with two perpendicular orientations that allow parallel and perpendicular polarized scattered light go through. Then, two Raman spectra are recorded: one records the intensity I_{VV} (the subscripts stand for Vertical, Vertical) the other records I_{VH} (the subscripts stand for Vertical, Horizontal). The **depolarization ratio** of a Raman band is given by

$$\rho_p = \frac{\int I_{VH}(\tilde{\nu})d\tilde{\nu}}{\int I_{VV}(\tilde{\nu})d\tilde{\nu}} \tag{4.56}$$

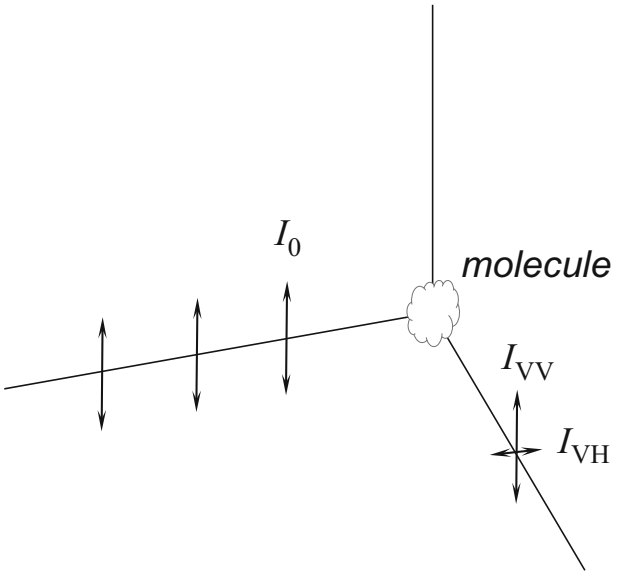


Fig. 4.20 Geometry for observing the depolarization ratio in a Raman spectrum

where the integrals extend over each band of the Raman spectrum. (Fig. 4.20).

Theoretical considerations enable one to conclude that the depolarization ratio of a Raman band lies in the range

$$0 \leq \rho_p \leq 3/4 \tag{4.57}$$

A Raman band is said to be *depolarized* when $\rho_p = 3/4$, *polarized* when $0 < \rho_p < ¾$, and *completely polarized* when $\rho_p = 0$. When the symmetry of a molecule is not preserved during a particular vibration (non-totally symmetric vibration), the depolarization ratio takes the value 3/4. For totally symmetric vibrations, ρ_p is smaller than 3/4.

Consider now the tetrahedral molecules shown in Fig. 4.21 belonging to the point groups C_{3v}, C_{2v}, C_s, and C_1. The following table shows the symmetry species for the vibrations of these molecules. Note that the numbers of polarized bands, that is, of totally symmetric vibrations, are different for these point groups, namely, 3 for C_{3v}, 4 for C_{2v}, 6 for C_s, and 9 for C_1. Thus, recording of the Raman spectra and measuring depolarization ratios enables one to assign each of these molecules to its point group (**E10**).

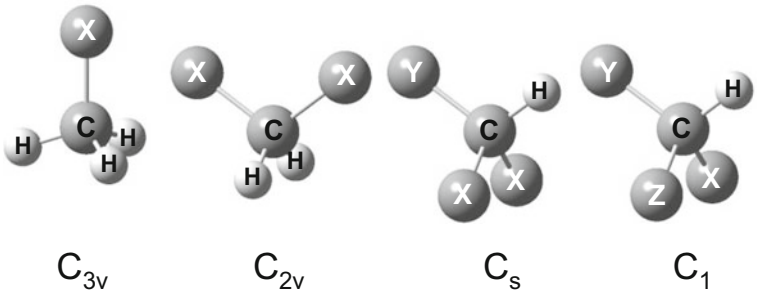


Fig. 4.21 Tetrahedral molecules belonging to different point groups; X, Y, and Z represent different halogen atoms

Symmetry species for vibrations of different point groups

Point group	Symmetry species
C_{3v}	$3A_1 + 3E$
C_{2v}	$4A_1 + A_2 + 2B_1 + 2B_2$
C_s	$6A' + 3A''$
C_1	$9A$

Mathematica Codes

M1. Classes of Symmetry Operations

In the following code, the symmetry operations of the C_{3v} group are represented by 3×3 matrices in the orthogonal basis $(x\ y\ z)^T$, with the matrices for σ_{v2} and σ_{v3} being obtained by similarity transformations on σ_{v1}, with X equal to C_3 and C_3^2, and the concept of class is illustrated. Since $(x\ y\ z)^T$ is an orthogonal system of axes, the inverses of matrices are obtained simply by transposing them. The *Mathematica* function `Transpose` transposes a matrix, and `If[condition,t,f]` gives `t` if `condition` evaluates to `True`, and `f` if it evaluates to `False`.

```
e = {{1,0,0}, {0,1,0}, {0,0,1}};
c3 = {{Cos[2Pi/3], -Sin[2Pi/3], 0}, {Sin[2Pi/3], Cos[2Pi/3], 0}, {0, 0, 1}};
c32 = {{Cos[4Pi/3], -Sin[4Pi/3], 0}, {Sin[4Pi/3], Cos[4Pi/3], 0}, {0, 0, 1}};
sv1 = {{-1,0,0}, {0,1,0}, {0,0,1}};
sv2 = Transpose[c3].sv1.c3;
sv3 = Transpose[c32].sv1.c32;
z = {{0,0,0}, {0,0,0}, {0,0,0}};
Print["      E          C3              C3^2          σv1             σv2             σv3"]
{e//MatrixForm,c3//MatrixForm,c32//MatrixForm,sv1//MatrixForm,sv2//MatrixForm,
    sv3//MatrixForm}
If[Transpose[sv1].c32.sv1-c3==z,Print["C3=(σv1)^T.C3^2.σv1"],Print["false"]]
If[Transpose[sv2].c32.sv2-c3==z,Print["C3=(σv2)^T.C3^2.σv2"],Print["false"]]
If[Transpose[sv3].c32.sv3-c3==z,Print["C3=(σv3)^T.C3^2.σv3"],Print["false"]]
If[Transpose[sv1].c3.sv1-c32==z,Print["C3^2=(σv1)^T.C3.σv1"],Print["false"]]
If[Transpose[sv2].c3.sv2-c32==z,Print["C3^2=(σv2)^T.C3.σv2"],Print["false"]]
If[Transpose[sv3].c3.sv3-c32==z,Print["C3^2=(σv3)^T.C3.σv3"],Print["false"]]
If[Transpose[c3].sv3.c3-sv1==z,Print["σv1=(C3)^T.σv3.C3"],Print["false"]]
If[Transpose[c32].sv3.c32-sv2==z,Print["σv2=(C3^2)^T.σv3.C3^2"],Print["false"]]
If[Transpose[c3].sv2.c3-sv3==z,Print["σv3=(C3)^T.σv2.C3"],Print["false"]]
If[Transpose[c32].sv2.c32-sv1==z,Print["σv1=(C3^2)^T.σv2.C3^2"],Print["false"]]
```

```
      E          C3              C3^2          σv1             σv2             σv3
{{{1,0,0},{0,1,0},{0,0,1}}, {{-1/2,-Sqrt[3]/2,0},{Sqrt[3]/2,-1/2,0},{0,0,1}},
 {{-1/2,Sqrt[3]/2,0},{-Sqrt[3]/2,-1/2,0},{0,0,1}}, {{-1,0,0},{0,1,0},{0,0,1}},
 {{1/2,-Sqrt[3]/2,0},{-Sqrt[3]/2,-1/2,0},{0,0,1}}, {{1/2,Sqrt[3]/2,0},{Sqrt[3]/2,-1/2,0},{0,0,1}}}
C3=(σv1)^T.C3^2.σv1
C3=(σv2)^T.C3^2.σv2
C3=(σv3)^T.C3^2.σv3
C3^2=(σv1)^T.C3.σv1
C3^2=(σv2)^T.C3.σv2
C3^2=(σv3)^T.C3.σv3
σv1=(C3)^T.σv3.C3
σv2=(C3^2)^T.σv3.C3^2
σv3=(C3)^T.σv2.C3
σv1=(C3^2)^T.σv2.C3^2
```

Consider the similarity transformation $\sigma_{v1}^{\mathrm{T}}.C_3.\sigma_{v1}$ with C_3 being a function of θ. We begin by representing σ_{v1} and $C_3(\theta)$ by the following matrices,

$$\sigma_{v1} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad C_3(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{4.58}$$

calculate $C_3(\theta).\sigma_{v1}$,

$$C_3(\theta) \cdot \sigma_{v1} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -\cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{4.59}$$

and $\sigma_{v1}^{\mathrm{T}}.C_3(\theta).\sigma_{v1}$

$$(\sigma_{v1})^T \cdot C_3(\theta) \cdot \sigma_{v1} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -\cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{4.60}$$

We now assume that $\alpha = 0$ and $r = 1$ in (4.61) (see Fig. 4.12),

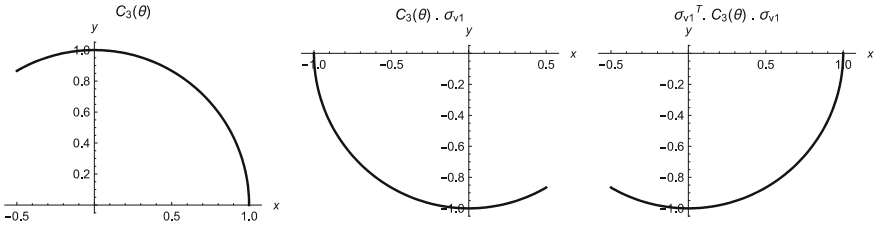
$$x = r\cos\ \alpha = 1 \quad y = r\sin\ \alpha = 0 \quad \Rightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{4.61}$$

and obtain

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = C_3(\theta)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = C_3(\theta) \cdot \sigma_{v1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\cos\theta \\ -\sin\theta \end{pmatrix}$$
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = (\sigma_{v1})^T \cdot C_3(\theta) \cdot \sigma_{v1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ -\sin\theta \end{pmatrix} \tag{4.62}$$

The following code plots $(x',y')^T$ for $C_3(\theta)$, $C_3(\theta).\sigma_{v1}$ and $\sigma_{v1}^{T}.C_3(\theta).\sigma_{v1}$ using `ParametricPlot`. This Mathematica function plots the (x', y')-coordinates as functions of a parameter (= external variable).

```
s=200; t=0.010;
aa=ParametricPlot[{Cos[u],Sin[u]},{u,0,2Pi/3},AxesLabel→{x,y},PlotLabel→"C3(θ)",
    PlotStyle→{Black,Thickness[t]},ImageSize→s];
bb=ParametricPlot[{-Cos[u],-Sin[u]},{u,0,2Pi/3},AxesLabel→{x,y},PlotLabel→"C3(θ).σv1",
    PlotStyle→{Black,Thickness[t]},ImageSize→s];
cc=ParametricPlot[{Cos[u],-Sin[u]},{u,0,2Pi/3},AxesLabel→{x,y},PlotLabel→"σv1T.C3(θ).σv1",
    PlotStyle→{Black,Thickness[t]},ImageSize→s];
Row[{aa,"  ",bb,"  ",cc," "}]
```



The first plot shows the arc of a circle of radius 1 ($r = 1$) drawn in the counterclockwise direction for θ from 0 to $2\pi/3$. The second plot shows the same rotation in the same direction (counterclockwise) *after* the initial point of the rotation was reflected on σ_{v1} (see Fig. 4.13). In turn, the third plot shows that the effect of multiplying on the left by σ_{v1}^{T} causes the curve to reflect by σ_{v1} (note that $\sigma_{v1}^{T} = \sigma_{v1}$, since the matrix representing σ_{v1} is diagonal), and this reflection amounts to a change of direction from counterclockwise to clockwise. On the whole, these

plots enable one to follow the $\sigma_{v1}^{T}.C_3(\theta).\sigma_{v1}$ orthogonal transformation "behind the scenes."

M2. Multiplication Table for the C_{2v} Group

The following *Mathematica* code builds the multiplication table for the group C_{2v}. To this end, the symmetry operations are represented by matrices in the basis $(x\ y\ z)^T$ [see (4.9)]. Note the way the command `If[condition,t,f]` is used, with the field `f` being used to include another `If`. For presenting the results, the function `Array` builds a 4×4 array called `a` and uses the command `Grid` to present a two-dimensional grid with items of size 2 by 2, where the elements of the array are included:

```
e = {{1,0,0}, {0,1,0}, {0,0,1}};
c2 = {{-1,0,0}, {0,-1,0}, {0,0,1}};
s1 = {{1,0,0}, {0,-1,0}, {0,0,1}};
s2 = {{-1,0,0}, {0,1,0}, {0,0,1}};
a1,1="E";a1,2="C2";a1,3="σxz";a1,4="σyz";
a2,1="C2";a3,1="σxz";a4,1="σyz";
a2,2="E"; a3,3="E"; a4,4="E";
If[c2.s1==s1,a2,3="σxz",If[c2.s1==s2,a2,3="σyz",f]];
If[c2.s2==s2,a2,4="σyz",If[c2.s2==s1,a2,4="σxz",f]];
If[s1.s2==s1,a3,4="σxz",If[s1.s2==s2,a3,4="σyz",If[s1.s2==c2,a3,4="C2",f]]];
a4,3="C2";a4,2="σxz";a3,2="σyz";
m=Grid[Array[a## &,{4,4}],Frame→All,ItemSize→{2,2}]
```

E	C_2	σ_{xz}	σ_{yz}
C_2	E	σ_{yz}	σ_{xz}
σ_{xz}	σ_{yz}	E	C_2
σ_{yz}	σ_{xz}	C_2	E

M3. Selection Rule for a Particle in a One-Dimensional Box

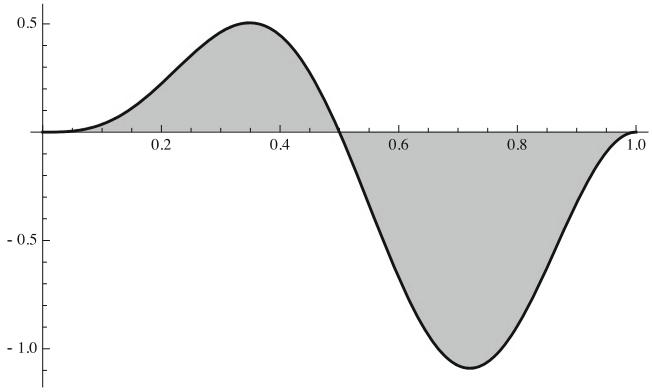
The following code calculates transition dipole moment integrals for a charged particle in a one-dimensional box. The first line defines the length of the box as 1, the second line presents the general expression for the one-dimensional box wave functions, the third line shows the values of the quantum numbers for the considered transitions (the list associated with i is outermost, that is, i changes more slowly than n), and the fourth line of code calculates the transition dipole moments for these transitions. The results indicate that for nonzero transition moments, the quantum number must change by an odd integer, ± 1, ± 3, …. The last line of code plots the integrand for the $1 \rightarrow 2$ transition, so that we can understand why the nonzero calculated transition dipole moments are negative. The option

`Filling→Axis` specifies that the area under the plotted curve should be filled to the axis. Note that the factor x of the integrand for the transition dipole moment weights more high values of x, thus giving rise to a negative area greater than the positive area and so to a negative integral.

```
d=1;
f[x_,n_]:= Sqrt[2/d] Sin[n Pi x/d]
Table[{n,n+i},{i,1,4},{n,1,3}]
Table[Integrate[f[x,n+i]*x*f[x,n],{x,0,d}],{i,1,4},{n,1,3}]
Plot[f[x,2]*x*f[x,1],{x,0,d},PlotStyle→Black,Filling→Axis]
```

```
{{{1, 2}, {2, 3}, {3, 4}}, {{1, 3}, {2, 4}, {3, 5}},
 {{1, 4}, {2, 5}, {3, 6}}, {{1, 5}, {2, 6}, {3, 7}}}
```

$$\left\{\left\{-\frac{16}{9\,\pi^2}, -\frac{48}{25\,\pi^2}, -\frac{96}{49\,\pi^2}\right\}, \{0, 0, 0\}, \left\{-\frac{32}{225\,\pi^2}, -\frac{80}{441\,\pi^2}, -\frac{16}{81\,\pi^2}\right\}, \{0, 0, 0\}\right\}$$

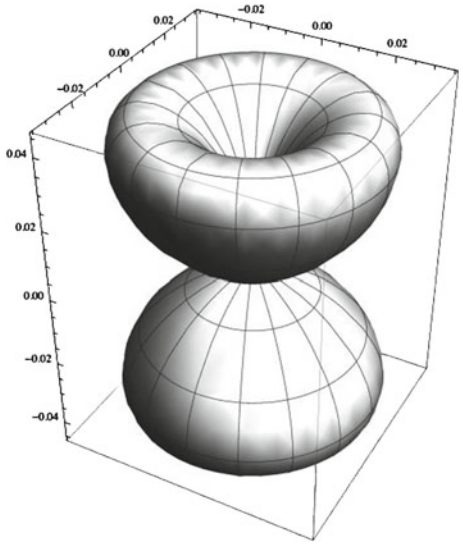


M4. Selection Rule for the Two-Particle Rigid Rotor

The following code calculates transition dipole moment integrals for the two-particle rigid rotor, whose eigenfunctions are spherical harmonics (see 3.§2). The first line shows the quantum numbers for the considered transitions, and the second line calculates the transition dipole moments and shows that allowed transitions have $\Delta J = \pm 1$ (J is the rotational quantum number). Since the spherical harmonics are functions of the angular coordinates, x should be substituted by $\cos\theta$, since $x = r\cos\theta$, and r is a constant. In addition, the integrand should include an additional $\sin\theta$ factor coming from the volume element, which is given by $\sin\theta\, d\theta\, d\phi$. The last line of code plots the integrand for the $0 \rightarrow 1$ J-transition. Note that the command for the plot is `SphericalPlot3D`, since the integrand is a function of angular coordinates θ and ϕ (`Plot3D` plots f as a function of x and y):

```
Table[{J,J+i},{i,1,4},{J,0,2}]
Table[Integrate[SphericalHarmonicY[J+i,0,θ,ϕ]*Cos[θ]*SphericalHarmonicY[J,0,θ,ϕ]*Sin[θ],
    {θ,0,Pi},{ϕ,0,2Pi}],{i,1,4},{J,0,2}]
SphericalPlot3D[SphericalHarmonicY[1,0,θ,ϕ]*Cos[θ]*SphericalHarmonicY[0,0,θ,ϕ]*Sin[θ],
    {θ,0,Pi},{ϕ,0,2Pi},PlotStyle→Gray,Lighting→"Neutral"]
```

```
{{{0, 1}, {1, 2}, {2, 3}}, {{0, 2}, {1, 3}, {2, 4}},
 {{0, 3}, {1, 4}, {2, 5}}, {{0, 4}, {1, 5}, {2, 6}}}
{{1/Sqrt[3], 2/Sqrt[15], 3/Sqrt[35]}, {0, 0, 0}, {0, 0, 0}, {0, 0, 0}}
```



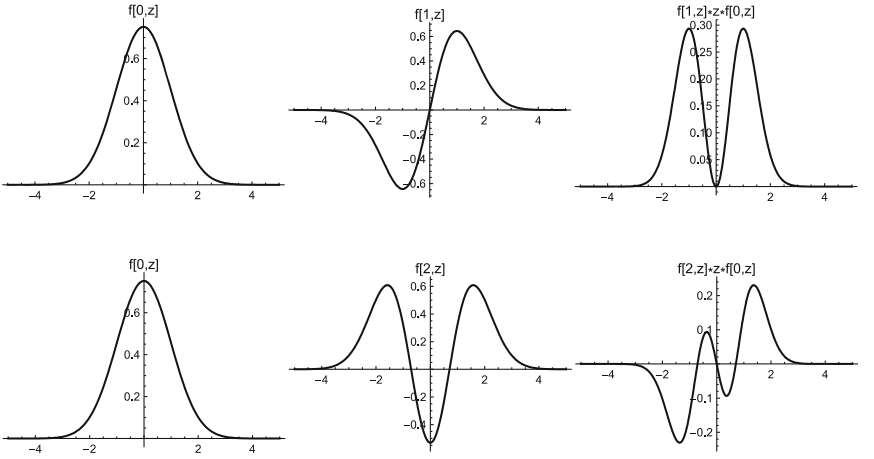
M5. Selection Rule for the Harmonic Oscillator

The following code calculates transition dipole moment integrals for the harmonic oscillator. The first line of code represents the eigenfunctions as functions of z ($= \sqrt{\alpha}\, x$) [see (3.62) and (3.63)], the second line shows the quantum numbers for the considered transitions, and the third line calculates the corresponding transition dipole moments. The results indicate that nonzero transition moments correspond to $\Delta n = \pm 1$, where n is the vibrational quantum number. The following lines of code plot the eigenfunctions and the integrands of the transition dipole moment integrals involved in one allowed transition (the $0 \rightarrow 1$ transition, where the integrand is an even function) and one forbidden transition (the $0 \rightarrow 2$ transition, where the integrand is an odd function):

```
f[n_,z_]:=(1/Pi)^(1/4)/Sqrt[2^n n!] E^(-z^2/2) HermiteH[n,z]
Table[{n,n+i},{i,1,4},{n,0,2}]
Table[Integrate[f[n+i,z]*z*f[n,z],{z,-Infinity,Infinity}],{i,1,4},{n,0,2}]
aa=Plot[f[0,z],{z,-5,5},PlotStyle→Black,PlotLabel→"f[0,z]",ImageSize→220];
bb=Plot[f[1,z],{z,-5,5},PlotStyle→Black,PlotLabel→"f[1,z]",ImageSize→220];
cc=Plot[f[1,z]*z*f[0,z],{z,-5,5},PlotStyle→Black,PlotLabel→"f[1,z]*z*f[0,z]",
    ImageSize→220];
dd=Plot[f[2,z],{z,-5,5},PlotStyle→Black,PlotLabel→"f[2,z]",ImageSize→220];
ee=Plot[f[2,z]*z*f[0,z],{z,-5,5},PlotStyle→Black,PlotLabel→"f[2,z]*z*f[0,z]",
    ImageSize→220];
Row[{aa,bb,cc}," "]
Row[{aa,dd,ee}," "]
```

```
{{{0, 1}, {1, 2}, {2, 3}}, {{0, 2}, {1, 3}, {2, 4}},
 {{0, 3}, {1, 4}, {2, 5}}, {{0, 4}, {1, 5}, {2, 6}}}
```

$$\left\{\left\{\frac{1}{\sqrt{2}}, 1, \sqrt{\frac{3}{2}}\right\}, \{0, 0, 0\}, \{0, 0, 0\}, \{0, 0, 0\}\right\}$$



Glossary

Abelian group — A point group in which all products of symmetry operations are commutative

Allowed transition — A nonzero transition dipole moment or nonzero transition polarizability

Character — The sum of the diagonal elements of a matrix representing a symmetry operation

Character table — A two-dimensional table whose rows are irreducible representations of the point group and whose columns are classes of symmetry operations in the group. A character table also presents the functions x, y, z, x^2, y^2, z^2, xy, xz, yz,

	and the rotations R_x, R_y and R_z around the x-, y-, and z-axes distributed by the irreducible representations of the group they belong to
Class	A complete set of mutually conjugate symmetry operations, which are related to each other by similarity transformations $B = X^{-1} A X$ with respect to any other symmetry operation X of the group. Symmetry operations within the same class have the same character, since the similarity transformation preserves the trace
Conjugate elements	A and B are related by a similarity transformation $B = X^{-1} A X$ with respect to a third symmetry operation X of the group and have the following properties: (1) Every group element is conjugate to itself. (2) If B is conjugate to A, then A is conjugate to B. (3) If A is conjugate to B and C, then B and C are conjugate to each other
Even function	A function $f(x)$ such that $f(x) = f(-x)$. The integral of an even function from $-a$ to a is different from zero, because the integrals from $-a$ to 0 and from 0 to a are equal
Forbidden transition	A zero transition dipole moment or zero transition polarizability
Group	A set of elements (which may be symmetry operations of a particular molecule) that obey the following mathematical requirements: (1) the product of any two elements, including the product of an element and itself, is a member of the group; (2) the group contains an identity element; (3) each element has an inverse; (4) the product is associative, i.e., $(AB)C = A(BC)$
Irreducible representation	A representation that cannot be decomposed into representations of smaller dimension
Odd function	A function $f(x)$ such that $f(x) = -f(-x)$. The integral of an odd function from $-a$ to a is zero, because the integrals from $-a$ to 0 and from 0 to a differ only by their sign
Point group	A group of symmetry operations of a finite object such as a molecule that has at least one point unchanged under every symmetry operation of the group
Selection rule	Identifies the quantum states of a molecule that give rise to allowed transitions

Symmetry operation	Converts the initial configuration of an object into an indistinguishable final configuration. If the symmetry operation is not observed during its application, it is impossible to decide whether it was applied, because the initial and final configurations are indistinguishable. Each symmetry operation is carried out by a symmetry element that can be an axis of rotation, a plane of reflection, a point of inversion, or an axis of rotation combined with a perpendicular plane of reflection

Exercises

E1. What is the principal axis of symmetry in each of the following molecules: (a) $CHCl_3$; (b) PF_5; (c) SF_6; (d) $C_5H_5^-$. Give the point group of each of the following molecules: (e) CH_2Cl_2; (f) BF_3; (g) C_6H_6; (h) CO_2; (i) CH_4.
E2. Use *Mathematica* to show that the transpose of the matrix representation of C_n in the basis $(x, y, z)^T$ is equal to its inverse.
E3. Use *Mathematica* to find the matrix representation for S_4 in the basis $(x, y, z)^T$.
E4. Use *Mathematica* to show that the transpose of a product of 3×3 matrices is equal to the product of transposes of matrix factors in reverse order, $(A.B)^T = B^T.A^T$.
E5. Use *Mathematica* to show that the inverse of a product of 3×3 matrices is equal to the product of the inverses of the matrix factors in reverse order, $(A.B)^{-1} = B^{-1}.A^{-1}$.
E6. Prove that square matrices B and A related by a similarity transformation ($B = X^{-1}AX$) have the same sum of their diagonal elements.
E7. Show that

(a) every element of a group A is conjugate to itself;
(b) if A is conjugate to B, then B is conjugate to A;
(c) if A is conjugate to B and to C, then B is conjugate to C.

E8. Use *Mathematica* to find the multiplication table for the C_{3v} group. Is this an abelian group?
E9. Considering a molecule with a center of symmetry, then the symmetry species for infrared active vibrations are of u-type, whereas the symmetry species for Raman active vibrations are of g-type. Explain.
E10. Find the reducible representations for vibrations of tetrahedral molecules belonging to the point groups C_{3v}, C_{2v}, C_s, and C_1.

Further Reading

Cotton, F.A.: Chemical applications of group theory, 3rd edn. Wiley, New York (1990)
Kettle, S.F.A.: Symmetry and structure (readable group theory for chemists). Wiley, New York (1995)
Vincent, A.: Molecular symmetry and group theory, 2nd edn. Wiley, New York (2001)

Molecular Structure

5

Abstract

The first two sections of this chapter define electron probability density and electrostatic potential and illustrate these topics with the water molecule. The ways of apportioning electrons of a molecule to different regions of space are performed by population analysis methods. We illustrate the Mulliken population analysis method for the methane molecule and the natural bond orbitals method for methane, ethene, and ethyne molecules. Next, we present a typical potential energy surface with one first-order saddle point obtained using a combination of Morse functions and calculate intrinsic reaction coordinates for the isomerization reaction HCN → CNH and the symmetric reaction $Cl^- + H_3CCl \rightarrow ClCH_3 + Cl^-$. Potential energy profiles for the rotations around the C–C bonds of ethane and 1,2-dichloroethane are presented and discussed. In particular, the potential energy profiles for the staggered conformation of ethane and the synclinal and antiperiplanar conformations of 1,2-dichloroethane, combined with vibrational calculations, enable one to estimate the amplitudes of the corresponding torsional motions. The last section of this chapter comprises chiral molecules, the Cahn–Ingold–Prelog rules for distinguishing *R* and *S* enantiomers of carvone, the polarimeter, and the way optically active molecules interact with plane-polarized electromagnetic radiation. At the end of this chapter, the student can find several Mathematica codes (Natural Bond Orbitals for CH_4, Potential Energy Surface, Right and Left Helices, Optical Rotation) with detailed explanations of new commands, a glossary of important scientific terms, and a list of exercises, whose complete answers are given after the appendix.

J.J.C. Teixeira-Dias, *Molecular Physical Chemistry*,
DOI 10.1007/978-3-319-41093-7_5

5.1 Electron Probability Density

Consider a molecule with n electrons. The wave function for the ground state can be written as

$$\Psi(s_1, \ldots, s_n) \tag{5.1}$$

where each s_i represents the set of three spatial and one spin coordinates of electron i. The probability of finding electron 1 in volume $\mathrm{d}v_1$ for *any* position of the remaining electrons and *any* spin of all electrons is given by

$$\mathrm{d}p(\boldsymbol{r}_1) = \left[\int \ldots \int \Psi_0^*(s_1, \ldots, s_n)\Psi_0(s_1, \ldots, s_n)\mathrm{d}v_2 \ldots \mathrm{d}v_n d\sigma_1 \ldots \mathrm{d}\sigma_n\right]\mathrm{d}v_1 \tag{5.2}$$

where $\boldsymbol{r}_1 = (x_1, y_1, z_1)$ is the position vector in some reference system of Cartesian axes, the integral over the spatial coordinates runs from electrons 2 to n, and the integral over the spin coordinates runs from 1 to n. The electrons are indistinguishable (the labeling of electrons is arbitrary), and so the probability of finding *any* electron in a volume $\mathrm{d}v$ is n times (5.2),

$$\mathrm{d}p(\boldsymbol{r}) = \rho_0(\boldsymbol{r})\mathrm{d}v \tag{5.3}$$

where

$$\rho_0(\boldsymbol{r}) = n\left[\int \ldots \int \Psi_0^*(s_1, \ldots, s_n)\Psi_0(s_1, \ldots, s_n)\mathrm{d}v_2 \ldots \mathrm{d}v_n \mathrm{d}\sigma_1 \ldots \mathrm{d}\sigma_n\right] \tag{5.4}$$

is the **electron probability density** for the molecule's ground state.

The ground state electron probability density is a function of three independent variables that are the coordinates of each point in the three-dimensional space around the molecule's nuclei,

$$\rho_0 = \rho_0(x, y, z) \tag{5.5}$$

Having four variables (x, y, z, and the function value ρ_0), this function cannot be plotted as a whole, since the world we live in is a three-dimensional world. In order to reduce the number of variables by one, we take the set of all points with coordinates (x, y, z) that satisfy the equation

$$\rho_0(x, y, z) = c \tag{5.6}$$

where c is a constant called an **isovalue** (from Greek *isos "equal"*). The above equation can be solved to obtain the z value as a function of x and y, that is,

$$\rho_0(x, y, z) = c \Rightarrow z = f(x, y) \tag{5.7}$$

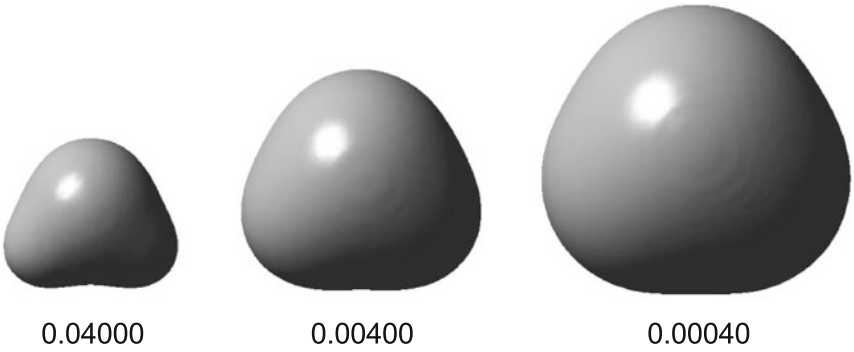


Fig. 5.1 Total electron probability density of the H_2O molecule represented by density surfaces with isovalues 0.04000, 0.00400, and 0.00040 (*Gaussian* 09 B3LYP/cc-pVTZ calculation with figure presented by *GaussView*)

Being a function of two variables, $f(x, y)$ represents a *surface* [recall that a function of one variable $y = g(x)$ represents a *curve*], and the surface derived from (5.6) is called an **isodensity surface**, because it is the set of points for which ρ_0 is equal to a specified isovalue.

Figure 5.1 presents three closed surfaces of the electron probability density for H_2O. Note that the surface with density isovalue 0.04000 is inside the surface with isovalue 0.00400, which, in turn, is totally contained within the surface whose isovalue is 0.00040. This relationship between surfaces with different isovalues resembles that of a set of Russian dolls: a smaller doll fits inside a larger one. The probability of finding any of the 10 H_2O electrons *within* the 0.00040 isodensity surface should be close to 100 %, and so this isodensity surface can be used to represent the overall shape of the molecule.

5.2 Electrostatic Potential

The force exerted by point charges q_1 and q_2 on each other is given by

$$F = \frac{q_1 q_2}{4\pi\varepsilon_0 r_{12}^2} \tag{5.8}$$

where r_{12} is the distance between the charges, and ε_0 is the permittivity of vacuum ($\varepsilon_0 = 8.854187817 \times 10^{-12}$ C^2 N^{-1} m^{-2}, see Appendix), and the potential energy of interaction is

$$E_p = \frac{q_1 q_2}{4\pi\varepsilon_0 r_{12}} \tag{5.9}$$

Consider now a simple charge distribution formed by two charges equal in magnitude and of opposite signs, in fixed positions (Fig. 5.2). Is the dominant electrostatic influence on point P_1 due to the positive or the negative charge? And on point P_2? Intuitively, the electrostatic influence of the negative charge on point P_1 should be greater than that of the positive charge on the same point, since the negative charge is closer to P_1 than the positive charge. In turn, at point P_2, the closer positive charge results in a dominant electrostatic effect. This suggests that the effect of each charge should be weighted by the inverse of the distance to the considered point. In addition, when a point is equally distant from two charges of the same sign and different magnitudes, the greater charge exerts the stronger electrostatic effect. Hence, apart from the factor $1/(4\pi\varepsilon_0)$, at P_1 the sum is negative, meaning that the dominant influence comes from the negative charge,

$$P_1(r_{1-} < r_{1+}) \quad \frac{q_-}{r_{1-}} + \frac{q_+}{r_{1+}} = -\frac{q}{r_{1-}} + \frac{q}{r_{1+}} < 0$$

where q is the absolute value of each charge. In turn, at P_2, the sum is positive, meaning that the dominant influence comes from the positive charge,

$$P_2(r_{2+} < r_{2-}) \quad \frac{q_-}{r_{2-}} + \frac{q_+}{r_{2+}} = -\frac{q}{r_{2-}} + \frac{q}{r_{2+}} > 0$$

The electrostatic influence of a distribution of charges on a particular point implies allocating a charge to that point, since charges interact with each other [see (5.9)]. However, the physical quantity to be determined cannot depend on a charge that does not belong to the original distribution of charges. In other words, the additional charge has to be so small that it does not significantly change the electric field created by the charge distribution. This infinitesimal positive charge at point P is called a **test charge** q_t. Let us now represent by $w(\infty \rightarrow P)$ the reversible work done by the electric field in providing the electrostatic force that transports q_t from infinity to the considered point P (Fig. 5.3). The electric field is defined in each point of space. Hence, the work done by the electric field to transport the test charge

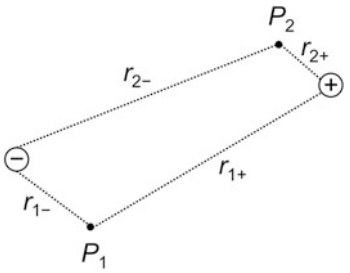


Fig. 5.2 The point P_1 "feels" a dominant electrostatic influence from the nearer negative charge, whereas P_2 "feels" a dominant electrostatic influence from the nearer positive charge

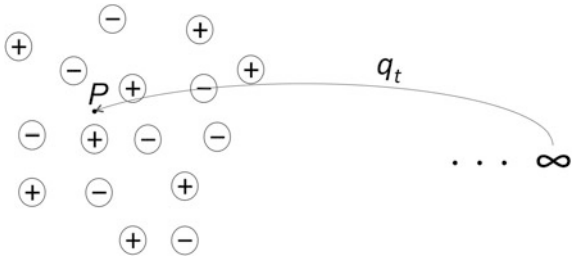


Fig. 5.3 The electrostatic potential at point P is equal to the reversible work done by the electric field in the transport of an infinitesimal test charge from infinity to point P, divided by the test charge

between two points in space depends only on these points, not on the path followed between them (the electrostatic force is a conservative force). In order to obtain a function defined at each point in space that does not depend on the test charge q_t, one divides $w(\infty \rightarrow P)$ by q_t, thus obtaining the **electrostatic potential** defined by

$$\phi(P) \equiv \frac{w(\infty \rightarrow P)}{q_t} \tag{5.10}$$

The electrostatic potential energy of q_t at infinity is null, because at infinity, q_t does not interact significantly with the charges of the distribution of charges.

The electrostatic potential difference between points A and B is given by the integral of the electric field over a path between these points,

$$\phi_B - \phi_A = -\int_A^B \mathcal{E}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} \tag{5.11}$$

where $\mathcal{E}$ is the electric field vector, whose magnitude is the ratio of the electrostatic force exerted on the test charge q_t and this charge, and $\mathrm{d}\boldsymbol{r}$ is the infinitesimal displacement vector along the path (A and B are the lower and upper limits of the integral). The electric field belongs to the system. When the electric field vector and the infinitesimal displacement vector have the same direction, $\mathcal{E}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} > 0$, the system does work on the exterior, thus releasing energy, and so the work is negative. In turn, when $\mathcal{E}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} < 0$, the exterior does work on the system, and the work is positive. When A is infinity (the test charge at infinity does not interact with the charge distribution, and so the electric potential at infinity is zero, $\phi_\infty = 0$), and B is the point P, the electrostatic potential ϕ_P due to a point charge is given by

$$\phi_P - \phi_\infty = -\int_\infty^P \mathcal{E}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} = -\frac{q}{4\pi\varepsilon_0}\int_\infty^r \frac{1}{r^2}\mathrm{d}r = \frac{q}{4\pi\varepsilon_0 r} \tag{5.12}$$

The generalization of (5.12) for a charge distribution is straightforward, since the electrostatic potential is an additive function, that is, the electrostatic potential of a charge distribution is the sum of the electrostatic potentials derived from all charges. In atomic units, where the unit of charge is e (the elementary charge), the unit of length is a_0 (the Bohr radius), and $4\pi\varepsilon_0 = 1$ (see Appendix), we can write

$$\phi(P) = \sum_i \frac{q_i}{r_{iP}} \tag{5.13}$$

where the charge q_i labeled by the subscript i includes not only its magnitude, but also its sign, and r_{iP} stands for the distance between the charge i and the point P where the electrostatic potential $\phi(P)$ is evaluated.

Consider now a point on the 0.00040 isodensity surface of H_2O (Fig. 5.1). The electrostatic potential of a molecule is the sum of two main contributions, one positive, given by the electrostatic potential produced by the nuclei with atomic numbers Z_n, the other negative, given by the electrostatic potential resulting from the electron charge density defined at each point in space. The electron charge density is obtained by multiplying the electron probability density ρ_0 by the electron charge $-e$. Since the electron charge is -1 in atomic units, the electron charge density is given in these units by $-\rho_0$, and the electron charge in the volume element dv is $-\rho_0 dv$. Hence,

$$\phi(P) = \sum_n \frac{Z_n}{r_{nP}} - \int \frac{\rho_0(\boldsymbol{r})}{r_{eP}} dv \tag{5.14}$$

where the second term is a triple integral over $dx\, dy\, dz$ (= dv) (**E1**).

Figure 5.4 shows the electrostatic potential for the H_2O molecule mapped over the total density surface with isovalue equal to 0.00040 using rainbow colors. Dark red indicates the negative extreme value of the electrostatic potential. Dark blue is used when the electrostatic potential reaches the extreme positive value. Positive values of the electrostatic potential in a particular region of space indicate a dominant influence of the nuclei in that region, which is said to be **electrophilic** (from Greek *philos "loving"*). In turn, negative values of the electrostatic potential in a particular region of space point to a dominant influence of the electrons, and that region is said to be **nucleophilic**. An electrophilic region of a molecule has a tendency to attract and interact with a nucleophilic region of another molecule. Such an interaction may result in a chemical reaction. Electrostatic potential maps for the active sites of enzymes are of great value in designing new drugs.

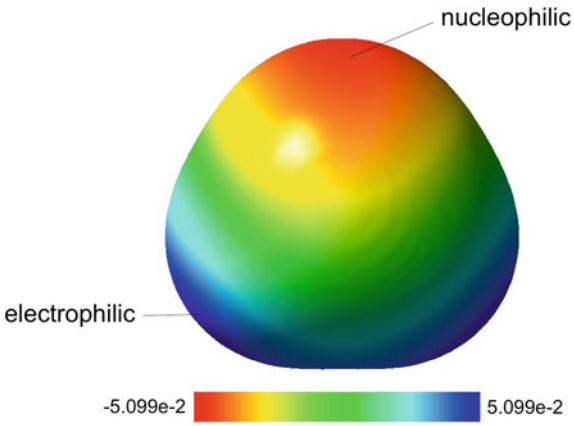


Fig. 5.4 Electrostatic potential for the H_2O molecule mapped over the total density surface with isovalue equal to 0.00040 (*Gaussian* 09 B3LYP/cc-pVTZ calculation with figure presented by *GaussView*)

5.3 Mulliken Population Analysis

The integral of the ground state electron density of a molecule over all space around the nuclei [see (5.4)] is equal to the total number of the molecule's electrons,

$$\int \rho_0(\boldsymbol{r})\mathrm{d}v = n \tag{5.15}$$

There are multiple ways of apportioning n to different regions of space in the molecule, such as atoms and overlap regions, or atoms and bonding regions. The methods of so doing are generally called **population analysis** and were developed initially by Mulliken, who assigned the total electron charge of a molecule to its atoms and overlap regions (Mulliken 1955).

5.3.1 Density Matrix

The ground state total electronic wave function for a closed-shell configuration is given by the following Slater determinant of molecular spin-orbitals,

$$\Psi_0 = (1, \ldots, n) = \left[\phi_1\bar{\phi}_1\phi_2\bar{\phi}_2 \ldots \phi_{n/2}\bar{\phi}_{n/2}\right] \tag{5.16}$$

where each column corresponds to a molecular spin-orbital, and the bar over ϕ_i indicates beta spin-orbitals (**E2**). The spatial factors of the molecular spin-orbitals are given by linear combinations of Slater-type atomic orbitals

$$\phi_i = \sum_{s=1}^{b} c_{si}\chi_s \tag{5.17}$$

Substitution of (5.16) and (5.17) in (5.4) gives

$$\rho_0(r) = 2\sum_{j=1}^{n/2} \phi_j^* \phi_j = 2\sum_{r=1}^{b}\sum_{s=1}^{b}\sum_{j=1}^{n/2} c_{rj}^* c_{sj}\chi_r^*\chi_s \tag{5.18}$$

and integration over all space leads to

$$\int \rho_0(r)\mathrm{d}v = 2\sum_{r=1}^{b}\sum_{s=1}^{b}\sum_{j=1}^{n/2} c_{rj}^* c_{sj} S_{rs} = \sum_{r=1}^{b}\sum_{s=1}^{b} P_{rs}S_{rs} = n \tag{5.19}$$

where P_{rs} represents the rs **density matrix** element and S_{rs} is the corresponding **overlap integral**,

$$P_{rs} = 2\sum_{j=1}^{n/2} c_{rj}^* c_{sj} \quad S_{rs} = \int \chi_r^*\chi_s \mathrm{d}v \tag{5.20}$$

Both the density and overlap matrices are symmetric matrices. For real coefficients,

$$P_{rs} = 2\sum_{j=1}^{n/2} c_{rj}c_{sj} = 2\sum_{j=1}^{n/2} c_{rj}c_{js}^{\mathrm{T}} = 2C_r C_s^{\mathrm{T}} \tag{5.21}$$

where C_r represents the r-row vector [$1 \times (n/2)$ vector] that multiplies the transpose of the s-row vector [$(n/2) \times 1$ vector]. The result of this vector product is a number, the P_{rs} element of the density matrix. In matrix notation, we can write

$$\boldsymbol{P} = 2\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}} \tag{5.22}$$

where the letters in boldface represent matrices.

The double summation over the atomic basis functions in the last equation of (5.19) includes diagonal terms for which $r = s$ and $S_{rr} = 1$ (the atomic basis functions are normalized) and off-diagonal terms for which $P_{rs}S_{rs} = P_{sr}S_{sr}$ (the lower and upper matrix triangles of the matrices $\boldsymbol{P}$ and $\boldsymbol{S}$ are equal, that is, these matrices are symmetric). Hence,

$$n = \sum_{r=1}^{b}\sum_{s=1}^{b} P_{rs}S_{rs} = \sum_{r=1}^{b} P_{rr} + 2\sum_{r>s}^{b}\sum_{s=1}^{b} P_{rs}S_{rs} \tag{5.23}$$

where the off-diagonal elements correspond to the matrix lower triangle (the row number is greater than the column number).

5.3.2 Minimal Basis Set Calculation for CH_4

In order to illustrate the Mulliken population analysis method without having to deal with extremely large matrices, we turn to a minimal basis set calculation for CH_4. This calculation was performed by the system of programs *Gaussian* 09, at the Hartree–Fock level, with the basis set STO-3G (HF/STO-3G calculation). The five doubly occupied lowest-energy molecular orbitals (CH_4 has 10 electrons) are mathematically expressed by the coefficients (eigenvectors) and corresponding energies (eigenvalues), as shown in the molecular orbital coefficient matrix. On application of the molecule's symmetry operations, these eigenvectors transform like two of the five irreducible representations of the T_d group, namely A_1 and T_2 (use the character table and see Fig. 5.5).

Molecular orbital coefficients

	A_1	A_1	T_2	T_2	T_2
Eigenvalues (eV)	−300.1	−24.8	−14.2	−14.2	−14.2
C 1*s*	0.99193	−0.22143	0.00000	0.00000	0.00000
C 2*s*	0.03836	0.62759	0.00000	0.00000	0.00000
C $2p_x$	0.00000	0.00000	0.00000	0.00000	0.57197
C $2p_y$	0.00000	0.00000	0.57197	0.00000	0.00000
C $2p_z$	0.00000	0.00000	0.00000	0.57197	0.00000
H $1s_1$	−0.00701	0.18057	0.30082	0.30082	0.30082
H $1s_2$	−0.00701	0.18057	−0.30082	0.30082	−0.30082
H $1s_3$	−0.00701	0.18057	0.30082	−0.30082	−0.30082
H $1s_4$	−0.00701	0.18057	−0.30082	−0.30082	0.30082

Character table for the T_d point group

	E	$8C_3$	$3C_2$	$6S_4$	$6\sigma_d$		
A_1	1	1	1	1	1		$x^2 + y^2 + z^2$
A_2	1	1	1	−1	−1		
E	2	−1	2	0	0		$(2z^2 - x^2 - y^2, x^2 - y^2)$
T_1	3	0	−1	1	−1	(R_x, R_y, R_z)	
T_2	3	0	−1	−1	1	(x, y, z)	(xy, xz, yz)

http://symmetry.jacobs-university.de

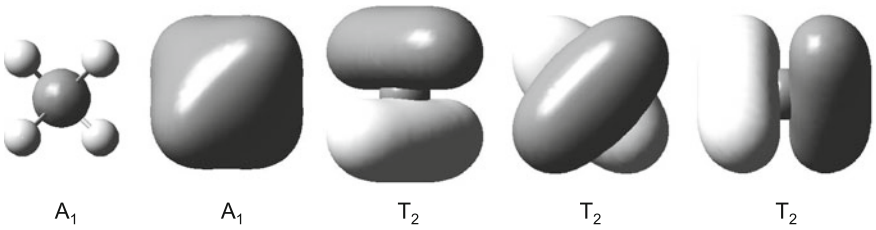


Fig. 5.5 Molecular orbitals for CH_4 represented by surfaces with isovalues equal to 0.06 and the corresponding symmetry species of the T_d point group. Lobes with different grades of *gray* correspond to different signs. HF/STO-3G Calculation performed by *Gaussian* 09, and molecular orbitals presented by *GaussView*

The first relevant feature of these orbitals is their delocalization over the whole molecule. In particular, all the doubly occupied orbitals have nonzero coefficients in each of the four hydrogen atoms' 1*s* orbitals. Note that the molecular orbital with the lowest energy is essentially an atomic orbital of the carbon atom core, since the largest coefficient refers to the carbon 1*s* orbital (the 2*p* orbitals have zero coefficients), and the coefficients for the hydrogen atoms' 1*s* orbitals are close to negligible. The molecular orbital with the second-lowest energy has its largest coefficient on the carbon 2*s* orbital. These two molecular orbitals transform like the A_1 irreducible representation of the T_d group, i.e., they are totally symmetric orbitals. The three molecular orbitals of highest energy are degenerate (they have the same energy eigenvalue). Each of these molecular orbitals has a C_2 axis that is coincident with each of the *x*-, *y*-, and *z*-axes and transforms like each of these coordinates. In addition, each of the C_2 symmetry axes coincides with an S_4 axis. The S_4 rotation–reflection symmetry operation includes a $2\pi/4$ rotation around the S_4 axis followed by a reflection in the plane perpendicular to the axis. Therefore, application of S_4 twice is equivalent to one C_2 symmetry operation. Note that the coefficients of degenerate molecular orbitals are not unique, since these can always be replaced by linear combinations that are orthogonal and correspond to the same energy eigenvalue.

The five molecular orbitals of methane are represented in Fig. 5.5 by surfaces whose isovalues are equal to 0.06. Considering the orientation and the different signs of the two lobes of each of the triple degenerate molecular orbitals, it can be concluded that the character of the S_4 rotation–reflection symmetry operation is equal to −1. Hence, these three molecular orbitals transform as the symmetry species T_2.

Density matrix

	C 1*s*	C 2*s*	C 2p_x	C 2p_y	C 2p_z	H 1s_1	H 1s_2	H 1s_3	H 1s_4
C 1*s*	2.06590								
C 2*s*	−0.20184	0.79067							
C 2p_x	0.00000	0.00000	0.65429						
C 2p_y	0.00000	0.00000	0.00000	0.65429					
C 2p_z	0.00000	0.00000	0.00000	0.00000	0.65429				
H 1s_1	−0.09387	0.22611	0.34412	0.34412	0.34412	0.60826			
H 1s_2	−0.09387	0.22611	−0.34412	−0.34412	0.34412	−0.11568	0.60826		
H 1s_3	−0.09387	0.22611	−0.34412	0.34412	−0.34412	−0.11568	−0.11568	0.60826	
H 1s_4	−0.09387	0.22611	0.34412	−0.34412	−0.34412	−0.11568	−0.11568	−0.11568	0.60826

The **density matrix** [matrix whose general elements are P_{rs}; see (5.21)] is a symmetric matrix, and so we present only the lower matrix triangle (**E3**). In addition, the elements of the **full Mulliken population analysis** matrix (see below) are given by $P_{rs}S_{rs}$ [see (5.19)], since they correspond to the product of the density

matrix elements P_{rs} and the corresponding overlap integrals S_{rs}. These can be obtained by dividing the Mulliken population matrix elements $P_{rs}S_{rs}$ by the corresponding density matrix elements P_{rs} (**E4**). Note that the diagonal elements for the density matrix and the full Mulliken population analysis are the same, since the diagonal overlap integrals are 1, $S_{rr} = 1$.

Full Mulliken population analysis matrix

	C $1s$	C $2s$	C $2p_x$	C $2p_y$	C $2p_z$	H $1s_1$	H $1s_2$	H $1s_3$	H $1s_4$
C $1s$	2.06590								
C $2s$	−0.05013	0.79067							
C $2p_x$	0.00000	0.00000	0.65429						
C $2p_y$	0.00000	0.00000	0.00000	0.65429					
C $2p_z$	0.00000	0.00000	0.00000	0.00000	0.65429				
H $1s_1$	−0.00594	0.11187	0.09331	0.09331	0.09331	0.60826			
H $1s_2$	−0.00594	0.11187	0.09331	0.09331	0.09331	−0.01993	0.60826		
H $1s_3$	−0.00594	0.11187	0.09331	0.09331	0.09331	−0.01993	−0.01993	0.60826	
H $1s_4$	−0.00594	0.11187	0.09331	0.09331	0.09331	−0.01993	−0.01993	−0.01993	0.60826

Gross orbital populations

C $1s$	C $2s$	C $2p_x$	C $2p_y$	C $2p_z$	H $1s_1$	H $1s_2$	H $1s_3$	H $1s_4$
1.99201	1.18802	1.02753	1.02753	1.02753	0.93433	0.93433	0.93433	0.93433

The sum of the elements of the Mulliken population matrix over one particular orbital, that is, over a row or a column, leads to the **gross orbital population** (see above), and the sum of the gross orbital populations over all atomic orbitals gives the total number of electrons [see (5.23)]. In turn, subtracting the sum of gross orbital populations belonging to a particular atom in the molecule from the total number of electrons in the isolated atom gives the **Mulliken atomic charges** (C −0.262620, H_1 0.065670, H_2 0.065670, H_3 0.065670, H_4 0.065670), whose sum (the total charge of the molecule) is zero. If the matrix elements of the Mulliken population analysis matrix corresponding to a specified atom are summed, one obtains the population **condensed** to that atom (see below) (**E5**)

Populations condensed to atoms (all electrons)

	1 C	2 H	3 H	4 H	5 H
1 C	4.71918				
2 H	0.38586	0.608263			
3 H	0.38586	−0.019929	0.608263		
4 H	0.38586	−0.019929	−0.019929	0.608263	
5 H	0.38586	−0.019929	−0.019929	−0.019929	0.608263

The strong dependence between the Mulliken population analysis results and the basis set used in the calculation enable us to conclude that this population analysis method is of reduced practical utility.

5.4 Natural Bond Orbitals

An alternative to the Mulliken population analysis consists in condensing the electron density of a molecule into the lowest few orbitals that, in localized 1-center and 2-center regions of the molecule, have the maximum electron occupancy and correspond to Lewis-like electron pairs localized in atoms and chemical bonds. In order to illustrate the fundamental concepts of the **natural bond orbital method**, we consider the methane molecule and take full advantage of the equivalence of the four C–H bonds.

A regular tetrahedral molecule like CH_4 can be included in a cube, as Fig. 5.6 shows: the carbon atom occupies the center of the cube, two of the hydrogen atoms occupy the vertices along a diagonal of one of the faces, and the other two hydrogen atoms are located on the perpendicular diagonal of the opposed cube face (Fig. 5.6a). The angle α to the center of a regular tetrahedron is called the **tetrahedral angle** (Fig. 5.6b). Simple trigonometry can show that the cosine of this angle is equal to $-1/3$ ($\alpha \approx 109.47°$), and Cartesian coordinates can be easily assigned to the vertices of a regular tetrahedron, as shown in Fig. 5.6c (**E6**).

5.4.1 Hybrid Atomic Orbitals

The electronic configuration of an isolated carbon atom is $1s^2\ 2s^2\ 2p^2$. Since all directions that emerge from the nucleus of an isolated atom are equivalent (the space around the nucleus is isotropic), the $2p$ electrons of an isolated carbon atom are uniformly distributed over the $2p_x$, $2p_y$, and $2p_z$ orbitals, which share the same electron density. Now consider the carbon atom of a CH_4 molecule and four equivalent independent linear combinations of the $2s$, $2p_x$, $2p_y$, and $2p_z$ orbitals directed to the hydrogen atoms on the vertices of a regular tetrahedron. These

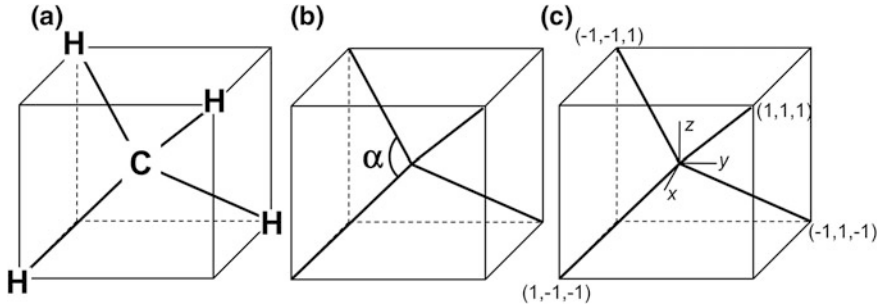


Fig. 5.6 **a** CH_4 inscribed in a cube. **b** Tetrahedral angle. **c** Cartesian coordinates for the vertices of a regular tetrahedron

orbitals can be made orthonormal and are called **hybrid atomic orbitals**. Having spherical symmetry, the 2*s* orbital participates in them with the same coefficient (we can begin by assigning 1, apart the normalization constant for the whole hybrid orbital). The subscript in each of the $2p_x$, $2p_y$, $2p_z$ orbitals indicates how these atomic orbitals transform under a symmetry operation. For example, $2p_x$ transforms like the x coordinate, that is, $2p_x = x\,f(r)$, where r is the distance from the point to the origin. Therefore, the $2p_x$, $2p_y$, and $2p_z$ orbitals act like unit vectors along the x-, y-, and z-axes, and their coefficients in the hybrid orbitals are given by the corresponding Cartesian coordinates of each vertex to which the hybrid orbital is directed (Fig. 5.6c), apart from the normalization constant.

The hybrid orbital directed to the (1, −1, −1) vertex is given by

$$N(2s + 2p_x - 2p_y - 2p_z) \tag{5.24}$$

where N is the normalization constant. Since the atomic orbitals are orthogonal and normalized, the normalization constant of the hybrid atomic orbital is given by the inverse of the square root of the sum of the squared orbital coefficients. In particular, for (5.24) we can write

$$N = 1/\sqrt{(+1)^2 + (+1)^2 + (-1)^2 + (-1)^2} = \frac{1}{2} \tag{5.25}$$

We now consider an sp^{λ} hybrid orbital

$$h_A = \frac{1}{\sqrt{1+\lambda}}\left(s + \sqrt{\lambda}p_A\right) \tag{5.26}$$

where p_A is a contracted p orbital (a linear combination of primitive p orbitals) directed to atom A. Since the s and p orbitals centered on the same atom are orthogonal, the squares of the coefficients of s and p_A,

$$f_s = \frac{1}{1+\lambda} \quad f_p = \frac{\lambda}{1+\lambda} \tag{5.27}$$

are the s and p **fractions**, and λ is the **hybridization parameter**, given by the ratio f_p/f_s (**E7**, **E8**, **E9**). Note that the sum of f_s and f_p is equal to 1.

Let us now consider the sp^{λ_A} and sp^{λ_B} hybrids directed to atoms A and B,

$$h_A = \frac{1}{\sqrt{1+\lambda_A}}\left(s + \sqrt{\lambda_A}p_A\right) \quad h_B = \frac{1}{\sqrt{1+\lambda_B}}\left(s + \sqrt{\lambda_B}p_B\right) \tag{5.28}$$

The orthogonality of h_A and h_B leads to

$$S_{p_A p_B} = \int p_A p_B \mathrm{d}v = -\frac{1}{\sqrt{\lambda_\mathrm{A}\lambda_\mathrm{B}}} \tag{5.29}$$

(**E10**). The scalar product of two unit vectors directed to atoms A and B is given by

$$\hat{d}_A.\hat{d}_B = \cos\omega \tag{5.30}$$

These unit vectors are aligned with p_A and p_B orbitals, which in turn behave like unit vectors whose scalar product is given by the overlap integral of p_A and p_B. Hence, we can equate the second members of (5.29) and (5.30) and write

$$-\frac{1}{\sqrt{\lambda_\mathrm{A}\lambda_\mathrm{B}}} = \cos\omega \tag{5.31}$$

an equality known as Coulson's theorem. Applying (5.31) to equivalent sp^λ hybrids leads to

$$\lambda = -\frac{1}{\cos\omega} \tag{5.32}$$

For $\lambda = 3$ (sp^3 hybrids), ω is equal to the tetrahedral angle (109.4712°), for $\lambda = 2$ (sp^2 hybrids), $\omega = 120°$, and for $\lambda = 1$ (sp hybrids), $\omega = 180º$.

If we now express the sp^λ hybrid in terms of the p_x, p_y, and p_z orbitals,

$$h = c_s s + c_x p_x + c_y p_y + c_z p_z \tag{5.33}$$

then λ is given by

$$\lambda = \frac{p\,\text{fraction}}{s\,\text{fraction}} = \frac{c_x^2 + c_y^2 + c_z^2}{c_s^2} \tag{5.34}$$

This expression shows that the hybridization parameter λ depends on the coefficients of s and p orbitals and so on the level and basis set of the calculation. In addition, (5.34) allows one to conclude that the λ value of an sp^λ atomic hybrid orbital cannot be restricted to integer values 3, 2, and 1. The integer values of λ are imposed by symmetry and correspond to four equivalent hybrid orbitals (T_d point group, example CH_4), three equivalent hybrid orbitals (D_{3h} point group, example BF_3) and two equivalent hybrid orbitals ($D_{\infty h}$ point group, example CO_2), respectively.

5.4.2 Natural Bond Orbitals for CH_4

Consider now the HF/STO-3G scan of the CH bond length R_{CH} in CH_4. Keeping the regular tetrahedron geometry, all HCH angles are equal to the tetrahedral angle, and the CH bonds have the same length. With these geometric constraints, there is only one independent variable, the CH bond length, R_{CH}. For large values of R_{CH}, the interactions between the carbon atomic orbitals and the hydrogen atomic orbitals are negligible, and so all orbitals reduce to those of isolated atoms. For R_{CH} smaller than 2.0 Å (this value may vary slightly according to the calculation level and the basis set of functions), the p fraction becomes equal to ¾, i.e., sp^3 hybrids are formed. Figure 5.7 shows the occupation of bond orbitals and the total molecular energy as functions of R_{CH}. At R_{CH} = 1.08 Å, the occupation of the orbitals attains a maximum, and the total molecular energy reaches a minimum. Molecular orbitals localized in the bonds that attain the maximum occupation allowed by the Pauli exclusion principle (n = 2) are called **natural bond orbitals** (NBOs). The carbon 1s orbital remains essentially a doubly occupied orbital of the carbon atom core and is thus called a **natural atomic orbital** (NAO). Each of the natural bond orbitals is a linear combination of one sp^3 hybrid on the carbon atom and the 1s orbital on the hydrogen atom to which the sp^3 hybrid is directed. In matrix form, we have

Orbital coefficients

C 1s	1.0000	0.0000	0.0000	0.0000	0.0000
C $(sp^3)_1$	0.0000	0.7253	0.0000	0.0000	0.0000
C $(sp^3)_2$	0.0000	0.0000	0.7253	0.0000	0.0000
C $(sp^3)_3$	0.0000	0.0000	0.0000	0.7253	0.0000
C $(sp^3)_4$	0.0000	0.0000	0.0000	0.0000	0.7253
H $1s_1$	0.0000	0.6885	0.0000	0.0000	0.0000
H $1s_2$	0.0000	0.0000	0.6885	0.0000	0.0000
H $1s_3$	0.0000	0.0000	0.0000	0.6885	0.0000
H $1s_4$	0.0000	0.0000	0.0000	0.0000	0.6885

The density matrix that corresponds to this coefficient matrix can be obtained by applying (5.22); it is given by

Density matrix

	C 1s	C $(sp^3)_1$	C $(sp^3)_2$	C $(sp^3)_3$	C $(sp^3)_4$	H $1s_1$	H $1s_2$	H $1s_3$	H $1s_4$
C 1s	2.0000								
C $(sp^3)_1$	0.0000	1.0521							
C $(sp^3)_2$	0.0000	0.0000	1.0521						
C $(sp^3)_3$	0.0000	0.0000	0.0000	1.0521					
C $(sp^3)_4$	0.0000	0.0000	0.0000	0.0000	1.0521				
H $1s_1$	0.0000	0.9987	0.0000	0.0000	0.0000	0.9481			
H $1s_2$	0.0000	0.0000	0.9987	0.0000	0.0000	0.0000	0.9481		
H $1s_3$	0.0000	0.0000	0.0000	0.9987	0.0000	0.0000	0.0000	0.9481	
H $1s_4$	0.0000	0.0000	0.0000	0.0000	0.9987	0.0000	0.0000	0.0000	0.9481

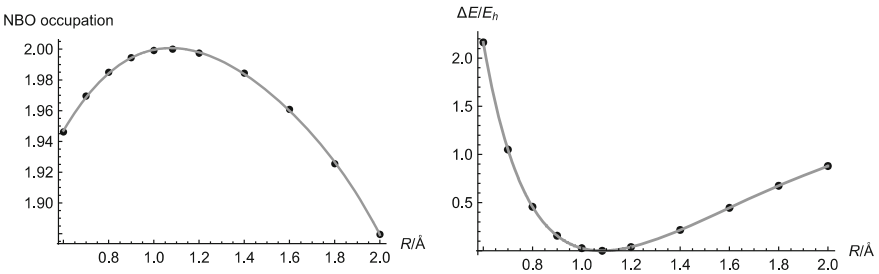


Fig. 5.7 Occupation of bond orbitals and total molecular energy of CH_4, as functions of R_{CH}. Natural bond orbitals (NBOs) occupation and total molecular energy as functions of R_{CH}, for regular tetrahedral arrangements of CH_4. Results from an HF/STO-3G calculation performed by *Gaussian* 09 and graphs obtained with *Mathematica*

The *Mathematica* code **M1** shows that the eigenvalues obtained by diagonalizing the above density matrix correspond to the maximum occupation number of the natural bond orbitals ($n = 2$), and the sum of the eigenvalues is equal to the total number of electrons of the molecule.

Figure 5.8 presents the natural bond orbitals for CH_4 represented by surfaces with isovalues equal to 0.06 (lobes with different shades of gray correspond to different signs). Unlike the canonical molecular orbitals that are delocalized over the CH_4 molecule and transform like the symmetry species of the T_d point group, the natural bond orbitals are equivalent orbitals localized in the C–H bonds, have maximum occupation ($n = 2$) and are in direct correspondence with Lewis dot structures.

5.4.3 Natural Bond Orbitals for $H_2C{=}CH_2$

We now turn to the ethene molecule (ethene is the systematic name for ethylene), $H_2C{=}CH_2$, where each carbon atom is bonded to three neighboring atoms: two hydrogen atoms and one carbon atom. The total number of valence electrons in the molecule is 12, four from each carbon atom, one from each hydrogen atom. The corresponding six electron pairs include four in the C–H bonds and two in the carbon–carbon double bond.

The natural bond orbitals for the valence electrons of the ethene molecule are represented by surfaces with isovalue equal to 0.06 and are shown in Fig. 5.9 by increasing order of energy. The lower row shows a natural bond orbital that has a nonzero electron density along the internuclear bond axis and results from a head-on overlap of atomic orbitals on both carbon atoms. It is called a **σ bond**. The middle row shows the C–H bonds that are equivalent by symmetry and have the same energy (are degenerate), and the natural bond orbital in the upper row is formed by two lobes, one above the plane of the nuclei, the other below, thus exhibiting zero electron density in the internuclear axis and in the plane of all the

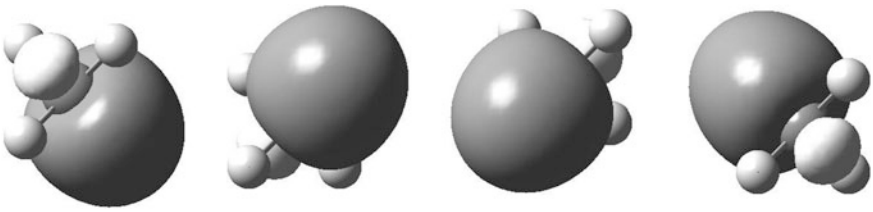


Fig. 5.8 Natural bond orbitals for CH_4 represented by surfaces with isovalues equal to 0.06, one per CH electron bond pair. HF/STO-3G Calculation by *Gaussian* 09, representation of molecule and natural bond orbitals presented by *GaussView*

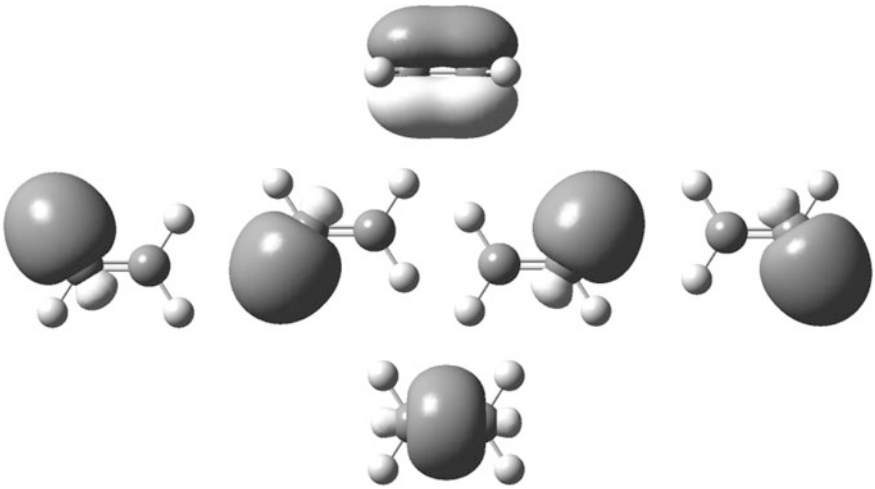


Fig. 5.9 Natural bond orbitals for $H_2C{=}CH_2$ illustrated by surfaces with isovalue equal to 0.06, one for each electron bond pair. They are ordered by increasing energy: one CC σ bond, four degenerate CH σ bonds, and one CC π bond (*Gaussian* 09 HF/STO-3G calculation; natural bond orbitals are drawn by *GaussView*)

nuclei (nodal plane). This bond results mainly from overlap of *p* orbitals with parallel axes that are perpendicular to the molecular plane and is called a **π bond**.

5.4.4 Natural Bond Orbitals for HC≡CH

In the ethyne molecule (ethyne is the systematic name for acetylene), HC≡CH, each carbon atom is bonded to two neighboring atoms: one carbon atom and one hydrogen atom. The ethyne molecule is linear. The total number of valence electrons in the ethyne molecule is 10, four from each carbon atom, one from each hydrogen atom. The corresponding five electron pairs include two electrons pairs for the carbon–hydrogen bonds and three electron pairs for the carbon–carbon triple

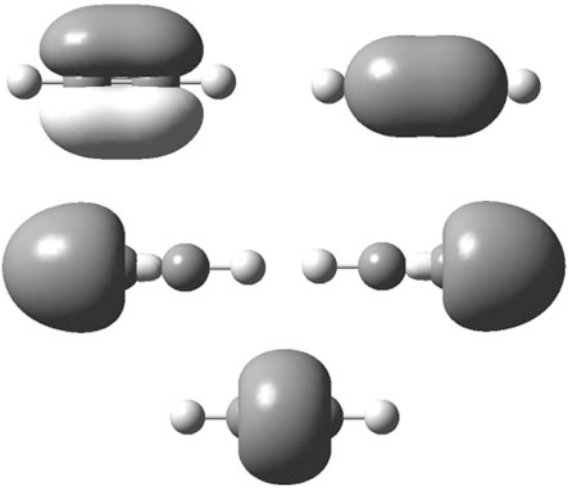


Fig. 5.10 Doubly occupied natural bond orbitals for the ethyne molecule are represented by surfaces with isovalue 0.06 and are shown by order of increasing energy: one CC σ bond, two degenerate CH σ bonds, two CC π bonds. HF/STO-3G Calculation performed by *Gaussian* 09, and natural bond orbitals visualized by *GaussView*

bond (one σ and two π bonds). The natural bond orbitals are presented in Fig. 5.10 by order of increasing energy, that is, E(σ C–C bond) < E(degenerate CH bonds) < E(degenerate π CC bonds).

5.4.5 CH Hybrids in CH_4, $H_2C{=}CH_2$ and HC≡CH

Considering now the CH hybrid orbitals in the methane, ethene, and ethyne molecules (find details of the calculations in Sect. 5.4.6), a clear trend can be observed, namely, the increasing *s* orbital fraction in the sp^{λ} hybrid orbitals, as we go from CH_4, to H_2CCH_2 and HCCH (see table below). The greater *s* fraction of the hybrid orbitals in the ethyne molecule allows negative charge to be held closer to the nucleus in the carbon atom, thus resulting in a greater acidic character of the hydrogen atoms in ethyne.

Carbon atom sp^{λ} hybrids and the corresponding *s* orbital fractions

Molecule	CH hybrids	*s* fraction	CC hybrids	*s* fraction
CH_4	sp^3	0.25	–	–
$H_2C{=}CH_2$	$sp^{2.37}$	0.30	$sp^{1.45}$	0.41
HC≡CH	$sp^{1.09}$	0.48	$sp^{0.91}$	0.52

5.4.6 Molecular Geometries and Electrostatic Potentials

The calculation for determining a molecular geometry consists in systematically varying the geometric variables until an energy minimum is reached. Nucleus–nucleus and electron–electron interactions involve particles of the same electric charge, thus being repulsive and corresponding to positive energy variations, that is, to energy variations that increase the total molecular energy. In turn, electron–nucleus interactions are attractive and involve negative energy variations, thus contributing to lower the molecular energy. When a balance between attractive and repulsive interactions is achieved for certain values of the geometric variables, the molecular energy reaches a stationary point, that is, a point with zero slope, and the calculation is called **geometry optimization**. In order to decide what type of stationary point (minimum or maximum) is obtained, a frequency calculation should be carried out using the stationary point geometry. If first derivatives increase in the vicinity of a stationary point (positive second-order derivatives), the stationary point corresponds to an energy minimum, and a frequency calculation will not encounter any imaginary vibrational frequency.

In order to calculate molecular geometries in close agreement with experimental values, we used a B3LYP calculation with a valence triple-zeta basis set (cc-pVTZ). The optimized geometric parameters for CH_4, C_2H_4, and C_2H_2 and the corresponding experimental values obtained from CCCBDB are shown in Fig. 5.11. The geometry optimization of CH_4 leads to a regular tetrahedral arrangement of the hydrogen atoms around the central carbon atom, that is, a T_d point group symmetry, and a CH bond length equal to 1.088 Å, in close agreement with the experimental value 1.087 Å (CCCBDB). The calculated geometric parameters of the ethene molecule compare very well with the corresponding experimental CCCBDB values, and those obtained for the ethyne molecule differ from the experimental CCCBDB values by orders of magnitude of 10^{-3} Å.

We used the optimized geometries to obtain the electrostatic potentials mapped over isodensity surfaces with isovalues equal to 0.000400 (Fig. 5.12). In agreement with the variation of the *s* orbital fraction values, the hydrogen atoms of ethyne

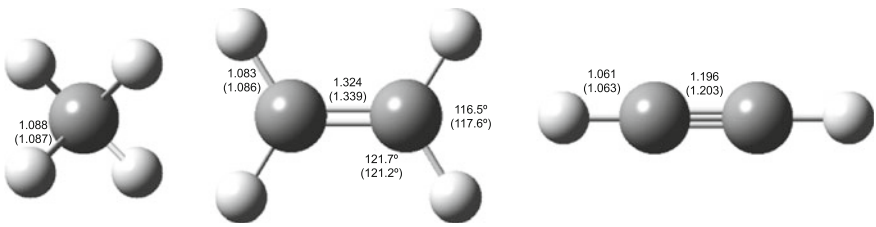


Fig. 5.11 Ball and bond representation of the ethane, ethene, and ethyne molecules with calculated versus experimental (CCCBDB) geometric parameters (bond lengths are in angstroms, bond angles in degrees, experimental values in parentheses). B3LYP/cc-pVTZ calculations performed by *Gaussian* 09 and molecular representations obtained with *GaussView*

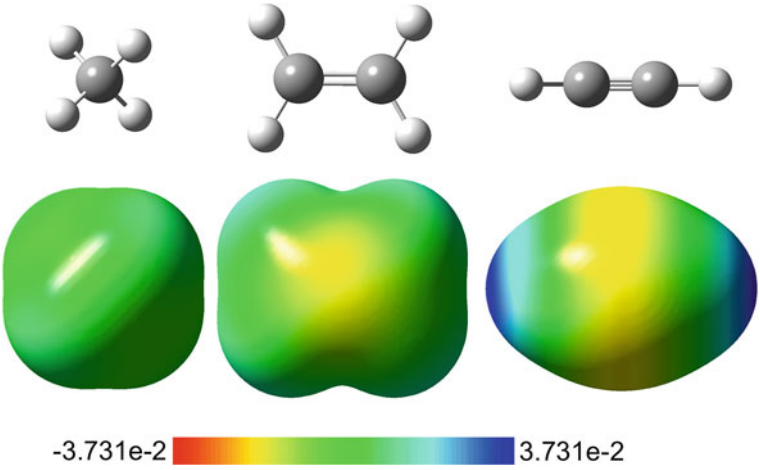


Fig. 5.12 Total density surfaces (isovalue = 0.000400) mapped with the corresponding electrostatic potentials, from *left to right*, for methane, ethene, and ethyne molecules. B3LYP/cc-pVTZ calculation performed by *Gaussian* 09, and representations of molecules and electrostatic potentials visualized by *GaussView*

exhibit a strong positive electrostatic potential, in contrast with the hydrogen atoms of ethene and methane.

5.5 Potential Energy Surfaces

Consider now the following hypothetical second-order reaction in gaseous phase:

$$A + BC \rightarrow AB + C \tag{5.35}$$

The system of three atoms, A, B, and C, can be described by three internal coordinates, namely, the internuclear AB distance r_{AB}, the internuclear BC distance r_{BC}, and the bond angle $\angle ABC$. For a defined angle, for example, $\angle ABC = \pi$ radians, the potential energy difference ΔU between the potential energy of the interacting three-atom system ($A^{\cdots}B^{\cdots}C$) and the sum of the energies of the three isolated atoms, A, B, and C, is a function of variables r_{AB} and r_{BC}, $\Delta U = f(r_{AB}, r_{BC})$. This function describes a surface called a **potential energy surface**.

From the mathematical point of view, a general potential energy surface can be characterized by its stationary points, i.e., by the points whose derivative is zero,

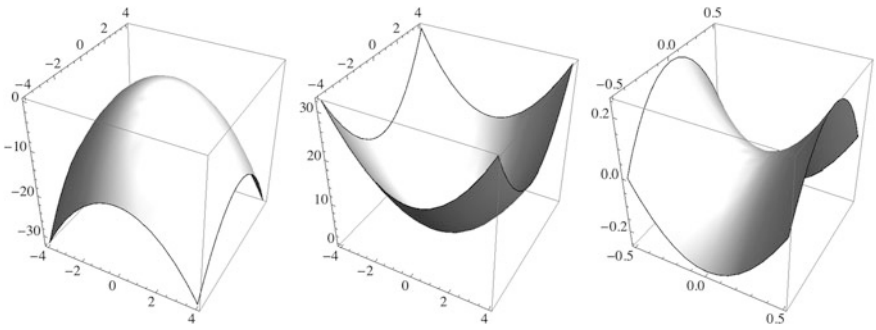


Fig. 5.13 Maximum, minimum, and saddle point at (0,0), for the functions $-x^2 - y^2$, $x^2 + y^2$ and $x^2 - y^2$. Figure obtained with *Mathematica*

$$\frac{\partial U}{\partial r_k} = 0 \quad k = 1, 2, \ldots, 3N - 6 \tag{5.36}$$

where r_k is each of the $3N - 6$ coordinates for a system with N atoms, and each stationary point can be classified as a maximum, a minimum, or a saddle point (Fig. 5.13). In a **first-order saddle point**, the function takes a maximum in one of the $3N - 6$ internal coordinates and minima in the remaining $3N - 7$ coordinates.

The potential energy surface of Fig. 5.14 corresponds to a linear triatomic system XHH and was obtained using Morse functions (see Baggott et al. 1988). It shows two valleys approximately parallel to the corresponding coordinate axes and an upper pass between them through one first-order saddle point. The reactant

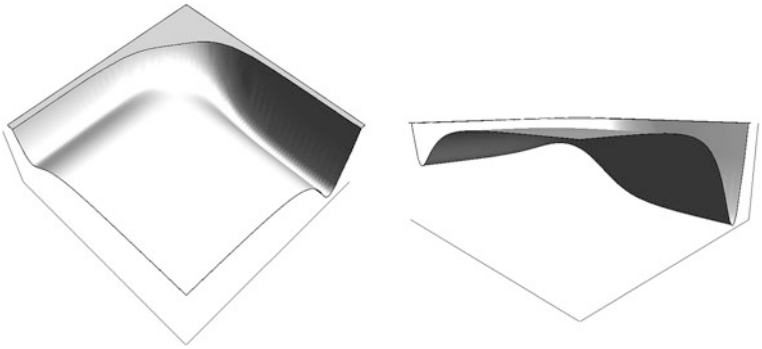


Fig. 5.14 Typical potential energy surface obtained using a combination of Morse functions. *Right* view shows the minimum energy path passing by the first-order saddle point. Figure obtained with *Mathematica*

channel and the product channel correspond to the chemical reaction $X + H_2 \rightarrow XH + H$ (see *Mathematica* code **M2**). The minimum-energy path converts reactants into products, passes through the first-order saddle point, and defines the **reaction coordinate**. The molecular state that corresponds to the maximum energy is the **transition state**.

5.5.1 Intrinsic Reaction Coordinate

When the reaction coordinate begins at the saddle point and goes down the hill toward the reactant channel (negative values of the reaction coordinate) and toward the product channel (positive values of the reaction coordinate), the obtained coordinate is called **intrinsic reaction coordinate**. Suggestively, the mathematical method used in the calculation for obtaining the intrinsic reaction coordinate is the **method of steepest descent**.

The linear triatomic molecules HCN (hydrogen cyanide) and HNC (hydrogen isocyanide) are isomers. Conversion of HCN into HNC implies hydrogen migration, where the transition state is given by a cyclic triangular structure (Fig. 5.15). The results presented in this figure show that the HCN → HNC reaction is endothermic.

Consider now the general reaction

$$Nu^- + RX \rightarrow RNu + X^- \tag{5.37}$$

where Nu^- stands for a nucleophile and X represents an electronegative atom bonded to a tetrahedral carbon atom in the R radical. A chemical reaction of this

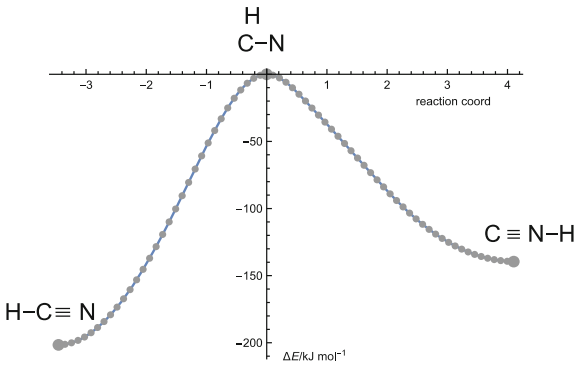


Fig. 5.15 Potential energy variation as a function of the intrinsic reaction coordinate for the isomerization reaction HCN → CNH (*Gaussian* 09 B3LYP/cc-pVTZ calculation). Figure obtained with *Mathematica*

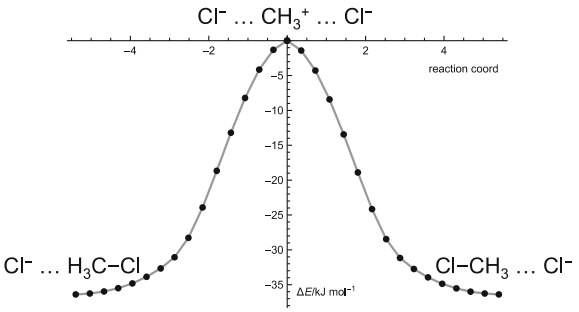


Fig. 5.16 Potential energy variation as a function of the intrinsic reaction coordinate for the reaction $Cl^- + H_3CCl \rightarrow ClCH_3 + Cl^-$ (*Gaussian* 09 B3LYP/aug-cc-pVTZ calculation). Figure drawn by *Mathematica*

type is called a **nucleophilic substitution reaction**. In fact, it is a substitution reaction (X is substituted by Nu) resulting from a nucleophilic attack of Nu^- on the carbon atom bonded to X:

$$Nu^- + C^{\delta+} - X^{\delta-} \tag{5.38}$$

If the nucleophilic substitution reaction occurs in a single step, the approach of Nu^- is *concerted* with the withdrawal of X^-. In this case, the transition state involves both reactant species Nu^- and RX, the reaction is bimolecular and named S_N2, and is said to occur in a **bimolecular concerted step**, with its rate equation given by

$$v = k[RX][Nu^-] \tag{5.39}$$

Consider now the bimolecular nucleophilic substitution reaction

$$Cl^- + H_3CCl \rightarrow ClCH_3 + Cl^- \tag{5.40}$$

whose potential energy variation as a function of the intrinsic reaction coordinate is shown in Fig. 5.16. In this academic example of a chemical reaction, products and reactants can be distinguished only by isotopic substitution and the conversion of reactants into products implies the **inversion** of the methyl group, a movement that recalls that of an umbrella suddenly struck by a blast of wind. In the transition state, the CH_3^+ carbocation presents a planar symmetrical configuration with the carbon atom equidistant from the chloride ions.

5.6 Molecular Conformations

5.6.1 Ethane

We now consider the ethane molecule, $H_3C–CH_3$, and the rotation of one of its methyl groups around the CC axis. The distances between the hydrogen atoms of the distinct methyl groups vary in the course of this methyl rotation (Fig. 5.17a) and so does the total molecular energy. In order to define the relative position of the methyl groups, four atoms bonded in sequence H(1)C(2)C(3)H(4) are considered, and the angle measured in the clockwise direction between the H(1)C(2)C(3) and C(2)C(3)H(4) planes is called a **dihedral angle** (Fig. 5.17b). Note that an angle between two planes is measured in the plane perpendicular to the common edge. Therefore, the dihedral angle for the ethane molecule is determined in the plane perpendicular to the CC bond axis (the plane of the paper in Fig. 5.17c). We consider the dihedral angle to vary in the range between 0° and 360°. A 60° dihedral angle for the atomic sequence H(1)C(2)C(3)H(4) corresponds to a 300° dihedral angle for the atomic sequence H(4)C(3)C(2)H(1).

The ethane molecule has one CC bond and six CH bonds. Around each tetrahedral center (each carbon atom), there are six bond angles (three CCH angles and three HCH angles), whose sum covers the whole space, thus being equal to 4π steradian (a constant). Hence, around each carbon atom there is one bond angle redundancy, meaning that there are five independent bond angle variables, not six. One additional coordinate is needed to define the relative position of the methyl groups, that being one HCCH dihedral angle. Altogether, we have seven bond lengths, plus 5 + 5 bond angles, plus one dihedral angle, making a total of 18 ($= 3N - 6$, where $N = 8$) geometric variables. In the potential energy plot of Fig. 5.18, the dihedral angle is the independent variable. Changing the dihedral angle will make the remaining $3N - 7 = 17$ variables adjust their values to each new value of the dihedral angle. One says that the dihedral angle scan is accompanied by **relaxation** of the molecular geometry (it is a **relaxed scan**).

The plot of the potential energy of the ethane molecule $H_3C–CH_3$ as a function of the dihedral angle HC–CH is called a **potential energy profile** for the rotation around the C–C bond (Fig. 5.18). As shown in this figure, for every initial value of the dihedral angle, the potential energy retakes the same value after a dihedral angle

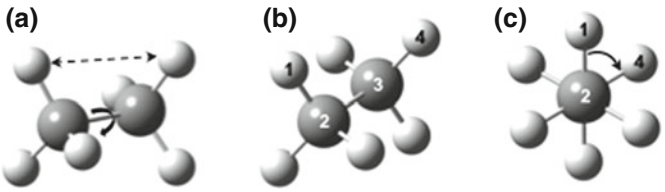


Fig. 5.17 **a** In the ethane molecule, methyl group rotation around the CC axis changes the distance between hydrogen atoms, thus causing the total molecular energy to vary. **b** Numbering of atoms involved in the dihedral angle. **c** The angle between 1-2-3 and 2-3-4 planes is measured in the plane perpendicular to the CC axis. Representation of molecules visualized by *GaussView*

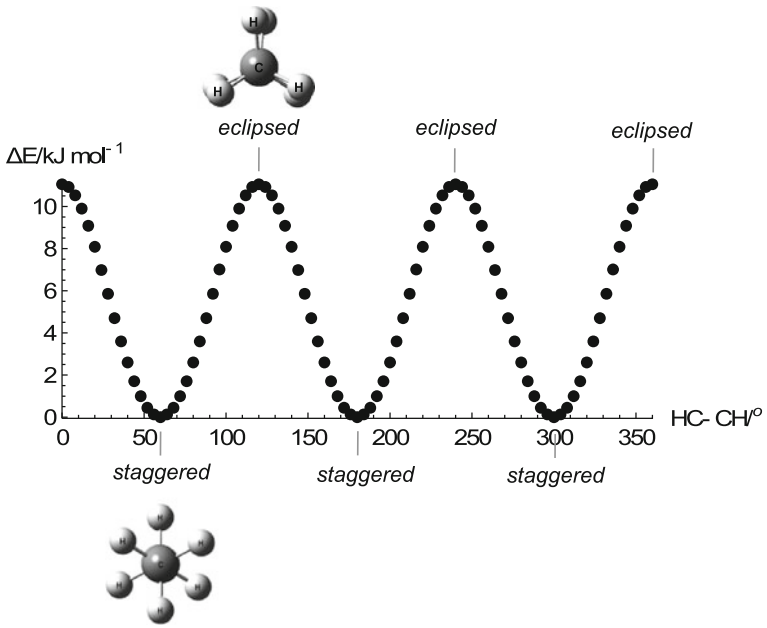


Fig. 5.18 Potential energy profile for the rotation of a methyl group in the ethane molecule. Staggered and eclipsed conformations correspond to stationary points, minimum, and maximum, respectively (*Gaussian* 09 B3LYP/cc-pVTZ calculation). Graph drawn with *Mathematica*

variation of 120° (= 360°/3). When the dihedral angle HC–CH is equal to 0° (= 360°), 120°, or 240°, the molecular energy reaches a maximum with respect to the dihedral angle. Looking at the molecule along the CC axis, the hydrogen atoms in the methyl group more distant from the observer are eclipsed by the hydrogen atoms of the methyl group closer to the observer, leading to a greater repulsion between hydrogen atoms of different methyl groups and, consequently, to a maximum in the potential energy profile. This is the **eclipsed** conformation. In turn, when the dihedral angle HC–CH is equal to 60°, 180°, or 300°, the molecule is at an energy minimum in the potential energy profile. This is the **staggered** conformation.

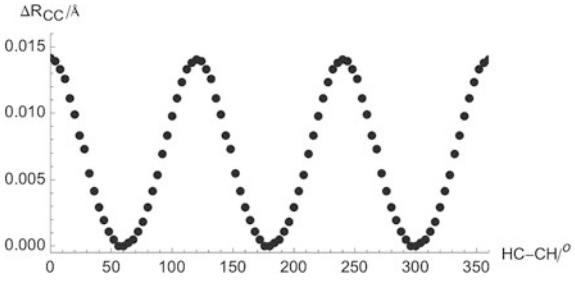


Fig. 5.19 CC bond length of the ethane molecule as a function of the dihedral angle (*Gaussian* 09 B3LYP/cc-pVTZ calculation; graph obtained with *Mathematica*)

The relaxation of geometric variables during the scan of the dihedral angle can be illustrated by the variation of the CC bond length. The observed variation of the CC bond length with the dihedral angle is *in phase* with the potential energy profile for the molecule, since the CC bond is along the threefold rotation axis of the molecule. The largest increase of the CC bond length attenuates the effect of CH interactions of distinct methyl groups in the eclipsed conformation (Fig. 5.19). In turn, the smallest values of the CC bond length correspond to the staggered conformation.

The ethane molecule has 18 normal modes of vibration ($N = 8$, $3N - 6 = 18$). Among these, the normal mode with the lowest frequency is the **torsional mode**, which is the out of phase oscillation of the methyl groups, whose atomic displacement vectors are shown in Fig. 5.20 for the staggered conformation, with the vectors applied to the hydrogen atoms of the same methyl group pointing in the same direction, either clockwise or counterclockwise, whereas the vectors applied to hydrogen atoms of different methyl groups point in opposite directions.

The calculated harmonic frequency for the torsional mode in the staggered conformation is 306 cm^{-1} (*Gaussian* 09 B3LYP/cc-pVTZ calculation). Figure 5.20 shows the potential energy curve with the minimum at the dihedral angle equal to 60° (staggered conformation), the zero-point energy level (level 0), and the first torsion level (level 1), which is 306 cm^{-1} above the zero-point level. The 0–1 transition corresponds to the **fundamental** transition, whereas the 0–2 transition is the first **overtone** (not shown). The amplitude of torsion in level 1 can be roughly estimated from Fig. 5.20 as ±30°. This means that each methyl group spans an arc of about 60°, and the total amplitude of motion relative to each other methyl group spans an arc of about 120° (the methyl groups rotate in opposite directions, one clockwise, the other counterclockwise). Hence, each methyl group probes practically all dihedral angles corresponding to different conformations around the staggered conformation.

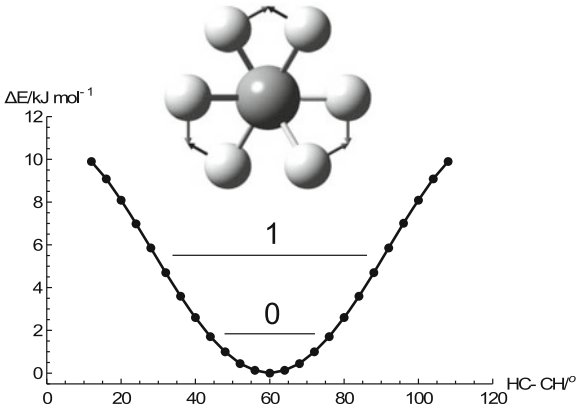


Fig. 5.20 Potential energy curve for the staggered conformation of ethane with harmonic energy levels for the zero-point energy and the $v = 1$ energy level of the torsional mode (*Gaussian* 09 B3LYP/cc-pVTZ calculation; molecule representation obtained with *GaussView*; graph obtained with *Mathematica*)

5.6.2 1,2-Dichloroethane

1,2-Dichloroethane is a toxic carcinogenic colorless liquid, mainly used in making chemicals involved in plastics, rubber, and synthetic textile fibers. Its production is primarily achieved by a catalyzed addition of chlorine to ethene. The relaxed scan of the dihedral angle ClC–CCl leads to the potential energy profile of Fig. 5.21. This is a symmetric profile with respect to the dihedral angle 180°, where the energy reaches a global minimum. In fact, from 180° to 360°, the curve is the reflected image of the curve from 0° to 180° (this symmetric behavior is also observed with respect to the dihedral angle 0°, where the energy reaches a global maximum). At the dihedral angle 0°, the conformation is **synperiplanar**, whereas at the dihedral angle 180°, is **antiperiplanar**. These conformations have reflection planes of symmetry that contain the nuclei of atoms in the ClCCCl dihedral sequence. The global energy minimum is obtained for the dihedral angle equal to 180°. From 0° to 180°, there are a local minimum (**synclinal** conformation) and a local maximum (**anticlinal** conformation). The energy difference between the global maximum and the global minimum in 1,2-dichloroethane is roughly four times the corresponding energy difference for the ethane molecule, an effect that can be ascribed to the Cl···Cl and Cl···H repulsive interactions involving the lone electron pairs of the bulky chlorine atoms.

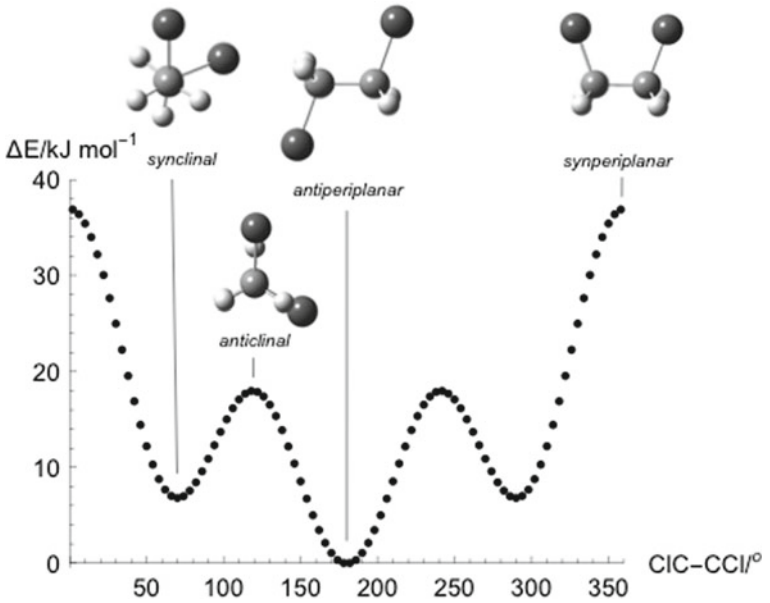


Fig. 5.21 Potential energy profile for the 1,2-dichloroethane molecule. Chlorine atoms are represented in *dark gray*. *Synclinal* and *antiperiplanar* conformations correspond to minima, *synperiplanar* and *anticlinal* conformations to maxima (*Gaussian* 09 B3LYP/cc-pVTZ calculation; molecule representation obtained with *GaussView*; graph obtained with *Mathematica*)

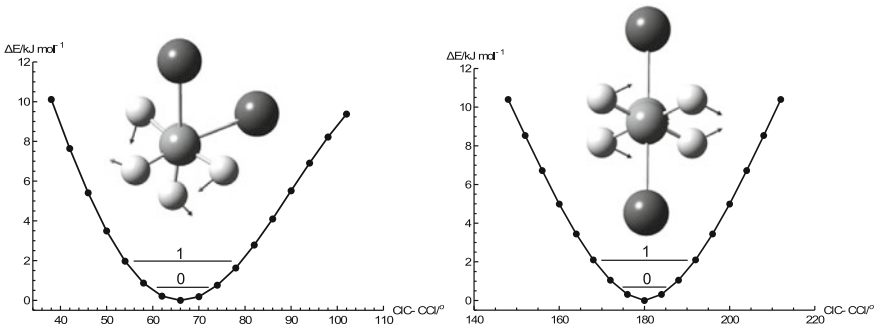


Fig. 5.22 Potential energy curves for the synclinal (*left*) and antiperiplanar conformations with the corresponding zero-point energy and $v = 1$ torsional levels and the displacement vectors for the corresponding vibrations (*Gaussian* 09 B3LYP/cc-pVTZ calculation; molecule representation obtained with *GaussView*; graph obtained with *Mathematica*)

Figure 5.22 shows the potential energy curves for the synclinal and antiperiplanar conformations, the displacement vectors for the corresponding torsional vibrations at the calculated values of 110 cm^{-1} for the synclinal conformation and 116 cm^{-1} for the antiperiplanar conformation, and the zero-point energy and $v = 1$ levels. The large volumes of the chlorine and carbon atoms hide their small displacement vectors. On each ClH_2 set of atoms, the displacement vectors point in the same direction, either clockwise or counterclockwise, whereas the vectors on different ClH_2 sets of atoms point in opposing directions. Note that the carbon atoms move slightly, so that the molecule's center of mass is kept still. Looking now at the amplitudes of the synclinal and antiperiplanar torsional motions in the $v = 1$ levels, we find that these can be roughly estimated from Fig. 5.22 as $\pm 12°$. The substitution of hydrogen in ethane by chlorine in 1,2-dichloroethane increases the reduced mass of the torsional oscillator, thus reducing the torsional frequency [see (3.55)] and, consequently, the amplitude of the torsional vibration. Unlike the methyl groups in ethane, the ClH_2C- groups of 1,2-dichloroethane have small amplitudes of torsional oscillations around their corresponding equilibrium positions. Thus, the synclinal and antiperiplanar conformations of 1,2-dichloroethane correspond to spectroscopically distinguishable **conformational isomers**, or **conformers**.

5.6.3 Boltzmann Distribution

Consider now a sample containing 1,2-dichloroethane in equilibrium at a specified temperature. What are the molecular populations of the antiperiplanar and synclinal

conformers? For a system in thermodynamic equilibrium, the population of the energy level ε_i (= the number of molecules with this particular energy) is given by the **Boltzmann distribution** (Ludwig Boltzmann, 1844–1906),

$$f_i = \frac{g_i \exp[-\varepsilon_i/(k_B T)]}{\Phi} \tag{5.41}$$

where

$$\Phi = \sum_k g_k \exp[-\varepsilon_k/(k_B T)] \tag{5.42}$$

T is the equilibrium temperature, k_B is the **Boltzmann constant** (= 1.38065×10^{-23} J K^{-1}; see Appendix) and g_i is the number of quantum states with energy ε_i, that is, the **degeneracy degree** (levels with a single quantum state have $g_i = 1$). Being a summation over k, Φ is independent of k and represents the total molecular population of all levels. Note that the range of each f_i is between 0 and 1, and the sum over all f_i equals 1. Hence, each f_i represents a **population fraction**. If the exponent of the exponential factor in (5.41) and (5.42) is multiplied and divided by Avogadro's constant, then the energy per molecule ε_i is substituted by the energy per mole, $E_i = N_A \varepsilon_i$ (N_A is Avogadro's constant), and the Boltzmann constant is replaced by the gas constant $R = N_A k_B$. Then we obtain

$$f_i = \frac{g_i \exp[-E_i/(RT)]}{\Phi} \tag{5.43}$$

where

$$\Phi = \sum_k g_k \exp[-E_k/(RT)] \tag{5.44}$$

For a system of two levels with degeneracy degrees g_1 and g_2, we can write

$$\Delta E_{21} = E_2 - E_1 = -RT \ln \frac{f_2/g_2}{f_1/g_1} \tag{5.45}$$

From (5.43) we can obtain the quotient f_i/g_i that appears in (5.45), for $i = 1$ and $i = 2$ (**E11**).

Vibrational spectroscopic studies in solution with a nonpolar solvent have to be carried out in order to obtain reliable experimental values concerning individual conformers. The first step consists in assigning vibrational bands to distinct conformers. This is not a simple task! Usually, it is accomplished with the help of relatively accurate frequency calculations for the considered conformers and of additional spectroscopic information that might be gathered. In the vibrational spectrum, one selects a pair of bands of the same vibrational mode in distinct conformers that allows a reliable determination of their intensities. The ratio of intensities of those bands equals the ratio of populations for the corresponding conformers.

5.7 Chiral Molecules

Molecules that are mirror images of one another have the same atoms and the same bonds, i.e., the same connectivity. Can we necessarily say that they are the *same* molecule with identical physical and chemical properties? A test that enables us to answer this question determines whether molecules that are mirror images of one other can be superimposed. Superposition is a virtual test, since molecules are impenetrable objects. Two molecules that are mirror images of each other and cannot be superimposed have a relationship analogous to a person's left and right hands.

Two chlorofluoromethane molecules, CH_2ClF, that are mirror images of each other can be superimposed (Fig. 5.23). Superposition implies a C_2 rotation of one of the molecules followed by its translation to match the positions of the atoms of the other molecule. The CH_2ClF molecules of this pair represent the same compound, with specific physical and chemical properties.

Replacement of the hydrogen atom of each CH_2ClF molecule away from the observer by one bromine atom results in two bromochlorofluoromethane molecules, CHBrClF (Fig. 5.23). Unlike CH_2ClF, the CHBrClF molecules exhibit a new feature: they are mirror images of each other but cannot be superimposed. No combined rotation and translation of one CHBrClF molecule of the pair can lead to superposition of both molecules. In fact, these molecules are to each other as the left and right hands of a person, and are thus called **chiral** objects (from Greek *kheir*, *hand*). Being isomers, they are also called **enantiomers** (from Greek *enantios* = *opposite*).

How are chirality and symmetry related? For a compound like CH_2ClF having a tetrahedral carbon atom as the central atom and not being chiral, we concluded that a C_2 rotation of the molecule followed by a translation leads to superposition with the other molecule of the object-to-image pair of molecules. In turn, a chiral molecule does not admit any S_n symmetry operation ($S_n = \sigma C_n$), including a reflection ($\sigma = S_1$) or an inversion ($i = S_2$).

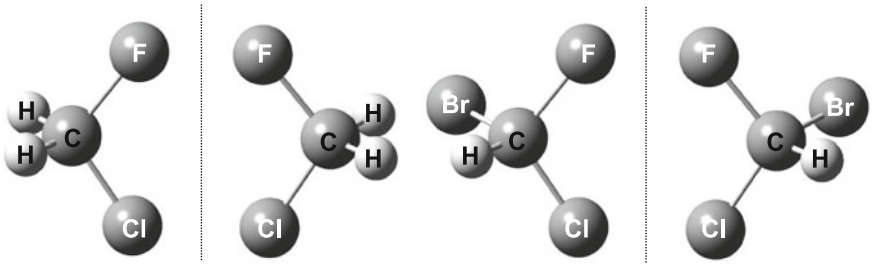


Fig. 5.23 *Left* Ball and bond representations of two superposable chlorofluoromethane molecules with an object-to-image plane mirror relationship to each other. *Right* Ball and bond representations of two nonsuperposable bromochlorofluoromethane molecules with an object-to-image relationship in a plane mirror to each other. Molecular representations obtained with *GaussView*

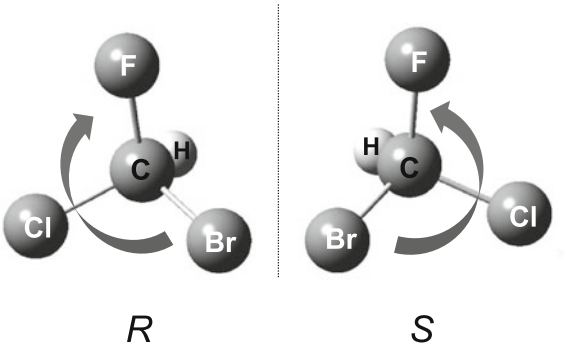


Fig. 5.24 *R* and *S* enantiomers of bromochlorofluoromethane

Molecules that are mirror images of each other and cannot be superimposed have a **chiral center**. The **Cahn–Ingold–Prelog rules** [Sidney Cahn (1899–1981, British chemist), Christopher Kelk Ingold (1893–1970, British chemist), and Vladimir Prelog (1906–1998, Croatian chemist), Nobel prize in chemistry in 1975 for advances in the stereochemistry of organic molecules and chemical reactions] assign different priorities to the atoms or groups bonded to a particular chiral center. These priorities follow the decreasing order of the atomic numbers. The atom or group with the lowest priority points away from the observer. If there is one hydrogen atom bonded to the chiral center, then the hydrogen atom is pointing away from the observer. If the decreasing priority of the atoms or groups bonded to the chiral center is in the clockwise direction, the chiral center is assigned an ***R*** (from Latin *rectus* = right) descriptor, whereas for the counterclockwise direction, the chiral center is assigned an ***S*** (from Latin *sinister* = left) descriptor (Fig. 5.24).

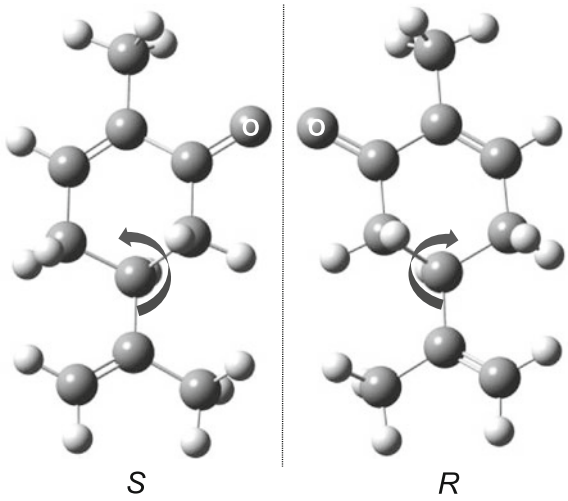


Fig. 5.25 *R* and *S* enantiomers of carvone

Carvone is an interesting chiral molecule: *R*-carvone smells like spearmint, *S*-carvone smells like caraway, a plant whose seeds are used as seasoning. The fact that enantiomers usually lead to widely different smells suggests that the olfactive receptors are chiral.

The *R* and *S* enantiomers of carvone are shown in Fig. 5.25. The chiral center is the tetrahedral carbon atom surrounded by curved arrows. Apart from the hydrogen atom pointing away from the observer, the atoms in the first layer of atoms bonded to the chiral center, CCC, score the same priority, since all of them are carbon atoms. The priority rules consider one double bond equivalent to two single bonds. Thus, applying priorities to the second layer of atoms, we obtain CCC > CHH=CHH. The tie between the last two members of the second layer is solved in the third layer, where the priority rules lead to OOC > CCH (see the following table). On the whole, applying the Cahn–Ingold–Prelog priority rules to carvone leads to the following sequence in order of decreasing priority: 1 = $-C(=CH_2)CH_3$, 2 = $-C(H_2)C=O$, 3 = $-C(H_2)CH=C$, 4 = H, as shown by the curved arrows around the chiral center in Fig. 5.25 (**E12**).

Layers of atoms around the chiral center of carvone and Cahn–Ingold–Prelog priorities

Layer	Atoms	Priorities
1	CCC	C=C=C
2	CCC, CHH, CHH	CCC > CHH=CHH
3	OOC, CCH	OOC > CCH

A **right-handed helix** follows a clockwise spiral path, whereas a **left-handed helix** follows a counterclockwise spiral path. Right- and left-handed helices that have the object-to-image relationship in a plane mirror cannot be superimposed (Fig. 5.26), and therefore represent chiral structures. Helicoid structures frequently appear in biological macromolecules such as nucleic acids and proteins. They also exist in quartz, where SiO_4 tetrahedra that share vertices form helicoid structures of a single kind (right- or left-handed helices). From the mathematical point of view, a right-handed helix can be obtained by combining a circle drawn in the clockwise

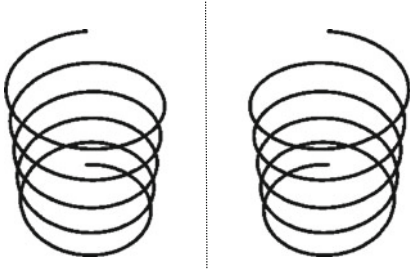


Fig. 5.26 *Left*- and *right*-handed helices having the object-to-image relationship in a plane mirror are not superimposable. Figure obtained with *Mathematica*

direction (a right-handed circle) with a perpendicular line segment. In a similar way, the left-handed helix results from a circle drawn in the counterclockwise direction (a left-handed circle) combined with a perpendicular line segment. The *Mathematica* code **M3** illustrates this point.

Before we find out how the enantiomers of a chiral molecule interact with electromagnetic radiation, we briefly consider a plane-polarized electromagnetic radiation. This consists of one electric field and one magnetic field with sinusoidal amplitudes that travel in phase in mutually perpendicular planes (Fig. 5.27). Usually, electrons in molecules interact mostly with the electric field. For that reason, the plane of polarization of the radiation is referred to the electric field plane.

The electric field component of plane-polarized electromagnetic radiation has two kinds of periodicity: *space* periodicity shows how the radiation progresses in a defined direction, and its characteristic parameter is the wavelength λ, whereas *time* periodicity shows how the electric field amplitude changes with time at a defined point along the propagation axis and is represented by the period τ. The inverse of period is the (linear) frequency v. The mathematical expression for a plane-polarized electric field polarized in the z direction and propagating along the x-axis is given by

$$\mathcal{E}_z(x,t) = \mathcal{E}_{z0}\cos(kx - \omega t + \theta) \tag{5.46}$$

where θ is the **phase**, k is the **angular wavenumber**, ω is the **angular frequency**,

$$k = \frac{2\pi}{\lambda} = 2\pi\tilde{\nu} \quad \omega = \frac{2\pi}{\tau} = 2\pi v \tag{5.47}$$

and $\mathcal{E}_{z0}$ is the electric field maximum amplitude, whose square of the absolute value is the energy carried by the wave.

Fig. 5.27 Plane-polarized electromagnetic radiation propagating along the x-axis. The electric field oscillates in the vertical plane. Figure obtained with *Mathematica*

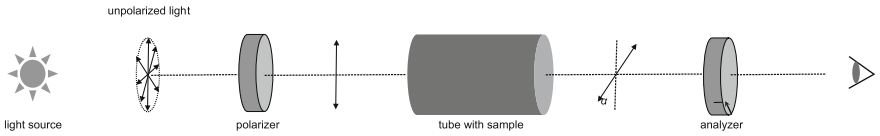


Fig. 5.28 Schematic setup for the main functions of a polarimeter

The enantiomers of a chiral molecule interact with electromagnetic radiation in symmetric but distinct ways: one of the enantiomers rotates the plane of polarization of light by a defined angle in the clockwise direction, whereas the other enantiomer rotates the plane of polarization by the same angle but in the counterclockwise direction.

When unpolarized radiation passes through a polarizer, plane-polarized electromagnetic radiation is obtained. If this radiation travels across a solution whose solute is a pure enantiomer, the polarization plane rotates by a specific angle and direction. For an observer to see light in the **polarimeter** (Fig. 5.28), the analyzer has to be rotated by the same angle of rotation, in the same direction, provided polarizer and analyzer were initially aligned. The solute molecule is said to be **optically active** or to exhibit **optical activity**. The magnitude and direction of the angle of rotation α depend on the chemical nature of the involved enantiomer, its concentration in solution c, and the interior length d of the sample tube, given by

$$\alpha = [\alpha]_{\lambda}^{T} cd \tag{5.48}$$

where $[\alpha]$ is called the **specific rotation**, i.e., the rotation for unit values of c and d. The temperature and wavelength dependencies of α are included in the specific rotation $[\alpha]$, measured for specified values of T and λ. The yellow sodium D line near the wavelength 589 nm is commonly used in measurements.

A plane-polarized electric field wave can be decomposed into left- and right-circularly polarized electric field waves (Fig. 5.29) that, in an optically active medium, have different refractive indices, thus propagating at different speeds.

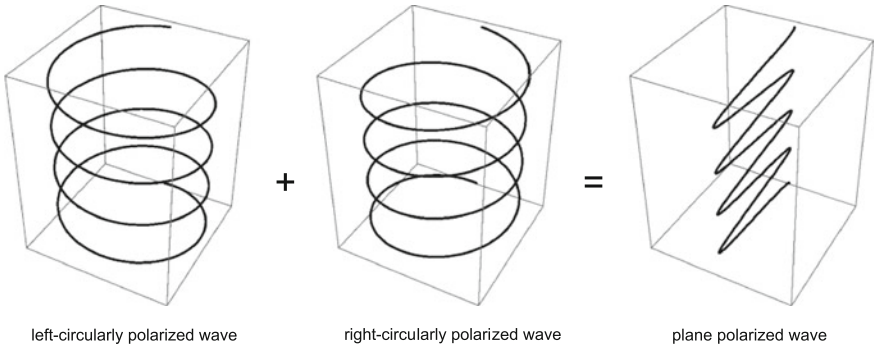


Fig. 5.29 A left-circularly polarized wave added to a right-circularly polarized wave gives a plane-polarized wave. Figure obtained with *Mathematica*

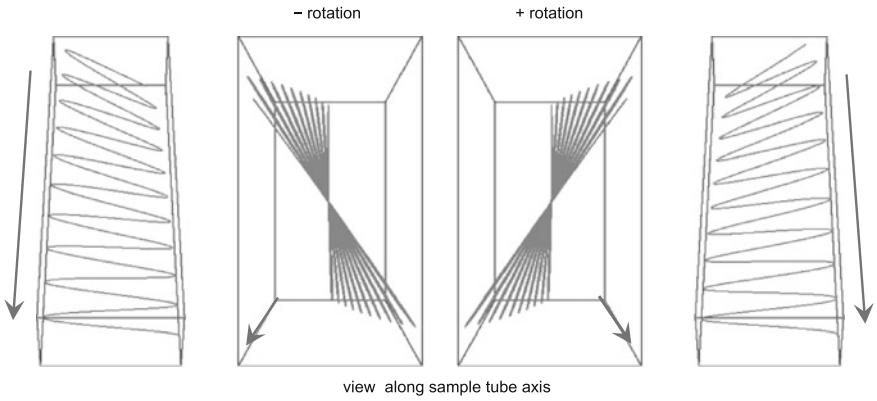


Fig. 5.30 On traversing an optically active medium, *left*- and *right*-circularly polarized waves change phase difference progressively, and the resulting wave gradually rotates. In a polarimeter, the views along the sample tube show the rotation angle for increasing lengths of the sample tube. Figure obtained with *Mathematica*

Hence, the phase relationship between left- and right-circularly polarized waves changes progressively, and the resulting wave gradually rotates (Fig. 5.30). The *Mathematica* code **M4** illustrates this point.

Mathematica Codes

M1. Natural Bond Orbitals for CH_4

Using the results of a *Gaussian* 09 HF/STO-3G calculation on CH_4, in particular the coefficient matrix that presents the five natural orbitals of CH_4 in terms of the 1*s* carbon orbital, four sp^3 hybrid atomic orbitals on carbon, and four 1*s* hydrogen atom orbitals, this *Mathematica* code shows that the eigenvalues obtained by diagonalizing the density matrix are 2 for the occupied orbitals (the maximum occupation number in accord with the Pauli exclusion principle) and 0 for the unoccupied orbitals, and the sum of the eigenvalues is equal to the total number of electrons of the molecule. The density matrix is obtained from the coefficient matrix using (5.22).

The *Mathematica* command `Chop` replaces real numbers whose absolute values are smaller than 10^{-4} by 0. The command `Transpose` transposes a matrix, and the commands `Eigenvalues` and `Eigenvectors` give the eigenvalues and eigenvectors of a square matrix.

```
Print["CH4 HF/STO-3G: (1 NAO + 4 NBOs) vs (C 1s + 4 C sp3 + 4 H 1s)"]
c = {{1, 0, 0, 0, 0}, {0, 0.7253, 0, 0, 0}, {0, 0, 0.7253, 0, 0},
  {0, 0, 0, 0.7253, 0}, {0, 0, 0, 0, 0.7253}, {0, 0.6885, 0, 0, 0},
  {0, 0, 0.6885, 0, 0}, {0, 0, 0, 0.6885, 0}, {0, 0, 0, 0, 0.6885}};
c//MatrixForm
Print["Density matrix (9x9)"]
d = 2 c.Transpose[c];
Chop[SetPrecision[d//MatrixForm,4],10^-4]
Print["Eigenvalues"]
Chop[SetPrecision[Eigenvalues[d]//MatrixForm, 4], 10^-4]
Print["Eigenvectors"]
vec = Transpose[Eigenvectors[d]];
(* in Mathematica, eigenvectors come as rows; Transpose converts them into columns *)
Chop[SetPrecision[vec//MatrixForm,4],10^-4]
```

```
CH4 HF/STO-3G: (1 NAO + 4 NBOs) vs (C 1s + 4 C sp3 + 4 H 1s)
```

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.7253 & 0 & 0 & 0 \\ 0 & 0 & 0.7253 & 0 & 0 \\ 0 & 0 & 0 & 0.7253 & 0 \\ 0 & 0 & 0 & 0 & 0.7253 \\ 0 & 0.6885 & 0 & 0 & 0 \\ 0 & 0 & 0.6885 & 0 & 0 \\ 0 & 0 & 0 & 0.6885 & 0 \\ 0 & 0 & 0 & 0 & 0.6885 \end{pmatrix}$$

```
Density matrix (9x9)
```

$$\begin{pmatrix} 2.000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.052 & 0 & 0 & 0 & 0.9987 & 0 & 0 & 0 \\ 0 & 0 & 1.052 & 0 & 0 & 0 & 0.9987 & 0 & 0 \\ 0 & 0 & 0 & 1.052 & 0 & 0 & 0 & 0.9987 & 0 \\ 0 & 0 & 0 & 0 & 1.052 & 0 & 0 & 0 & 0.9987 \\ 0 & 0.9987 & 0 & 0 & 0 & 0.9481 & 0 & 0 & 0 \\ 0 & 0 & 0.9987 & 0 & 0 & 0 & 0.9481 & 0 & 0 \\ 0 & 0 & 0 & 0.9987 & 0 & 0 & 0 & 0.9481 & 0 \\ 0 & 0 & 0 & 0 & 0.9987 & 0 & 0 & 0 & 0.9481 \end{pmatrix}$$

```
Eigenvalues
```

$$\begin{pmatrix} 2.000 \\ 2.000 \\ 2.000 \\ 2.000 \\ 2.000 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

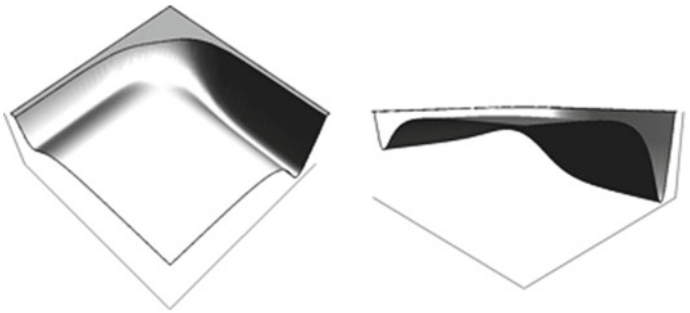
```
Eigenvectors
```

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1.000 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.7253 & 0 & 0 & -0.6885 & 0 & 0 & 0 \\ -0.7253 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.6885 \\ 0 & -0.7253 & 0 & 0 & 0 & 0 & 0.6885 & 0 & 0 \\ 0 & 0 & 0 & -0.7253 & 0 & 0 & 0 & -0.6885 & 0 \\ 0 & 0 & 0.6885 & 0 & 0 & 0.7253 & 0 & 0 & 0 \\ -0.6885 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.7253 \\ 0 & -0.6885 & 0 & 0 & 0 & 0 & -0.7253 & 0 & 0 \\ 0 & 0 & 0 & -0.6885 & 0 & 0 & 0 & 0.7253 & 0 \end{pmatrix}$$

M2. Potential Energy Surface

This potential energy surface is modeled for a linear triatomic system XHH as the sum of three Morse functions [see (3.68)], with one of them representing an interaction term (see Baggott et al. 1988). The channels linked by a first-order saddle point correspond to the chemical reaction $X + H_2 \rightarrow XH + H$. The dissociation energies for the reactant and product channels are equal to 20 and 50 arbitrary units of energy, and the force constants for the H_2 and XH diatomics are equal to 40 and 80 arbitrary units of energy per square length. The equilibrium point for each Morse function is defined at 1.2.

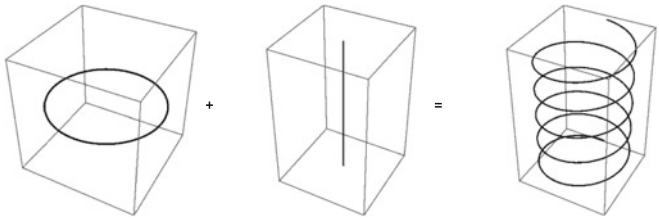
```
d1 = 50; k1 = 80; a1 = Sqrt[k1/(2*d1)];
d2 = 20; k2 = 40; a2 = Sqrt[k2/(2*d2)];
V1[r_, re_] := d1*(1 - Exp[-a1 (r - re)])^2;
V2[r_, re_] := d2*(1 - Exp[-a2 (r - re)])^2;
V12[r_, re_] := (d1 + d2)/2.*(1 - Exp[-(a1 + a2) (r - re)/2.])^2;
max = 10; min = 0.2;
Clear[f]
f[x_, y_] = V1[x, 1.2] + V2[y, 1.2] - 10 V12 [x + y, 1.2];
surface1 = Plot3D[f[x,y],{x,min,max},{y,min,max},PlotRange→{-330,-280},
    PlotStyle→Gray,Lighting→"Neutral",Mesh→None,PlotPoints→30,PlotStyle→Opacity[0.6],
    Boxed→False,ViewPoint→{1,1,2},Ticks→None,ImageSize→{250,250}];
surface2 = Plot3D[f[x,y],{x,min,max},{y,min,max},PlotRange→{-330,-280},PlotStyle→Gray,
    Lighting→"Neutral",Mesh→None,PlotPoints→30,PlotStyle→Opacity[0.6],
    Boxed→False,ViewPoint→{1,1,0.27},Ticks→None,ImageSize→{250,250}];
Row[{surface1,"  ",surface2}]
```



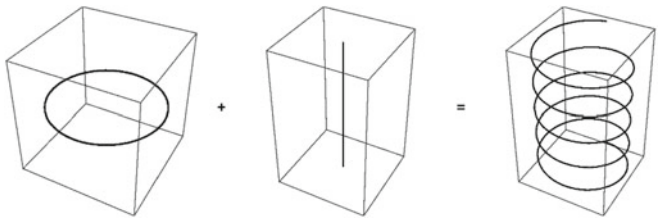
M3. Right- and Left-Handed Helices

The clockwise screwing motion of a right-handed helix moves the helix away from the observer. From the mathematical point of view, a right-handed helix can be obtained by combining a circle drawn in the clockwise direction with a perpendicular line segment. In a similar way, a left-handed helix can be seen as the result of combining a circle drawn in the counterclockwise direction with a perpendicular line segment. The next two *Mathematica* codes illustrate the mathematical ways of forming right- and left-handed helices. They use the `ParametricPlot3D` command, which produces a three-dimensional curve in which the x, y, and z component functions depend on a single parameter (an external variable) that can be interpreted as the time variable:

```
Rcircle = ParametricPlot3D[{-Sin[u],Cos[u],0},{u,0,2Pi},Axes→{False,False,False},
    Boxed→True,PlotStyle→{Thickness[0.01],Black},ImageSize→{250,250}];
line = ParametricPlot3D[{0,0,u/10},{u,0,10Pi},Axes→{False,False,False},Boxed→True,
    PlotStyle→{Thickness[0.01],Black},ImageSize→{250,250}];
Rhelix = ParametricPlot3D[{-Sin[u],Cos[u],u/10},{u,0,10Pi},Axes→{False,False,False},
    Boxed→True,PlotStyle→{Thickness[0.01],Black},ImageSize→{250,250}];
Row[{Rcircle,"  +  ",line,"  =  ",Rhelix}]
```

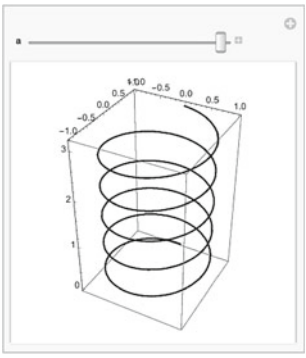


```
Lcircle = ParametricPlot3D[{Sin[u],Cos[u],0},{u,0,2Pi},Axes→{False,False,False},
    Boxed→True,PlotStyle→{Thickness[0.01],Black},ImageSize→{250,250}];
line = ParametricPlot3D[{0,0,u/10},{u,0,10Pi},Axes→{False,False,False},Boxed→True,
    PlotStyle→{Thickness[0.01],Black},ImageSize→{250,250}];
Lhelix = ParametricPlot3D[{Sin[u],Cos[u],u/10},{u,0,10Pi},Axes→{False,False,False},
    Boxed→True,PlotStyle→{Thickness[0.01],Black},ImageSize→{250,250}];
Row[{Lcircle,"  +  ",line,"  =  ",Lhelix}]
```



The following code uses the *Mathematica* command `Manipulate` to continuously vary the parameter range, thus visualizing the formation of the helix in the upward direction:

```
Manipulate[ParametricPlot3D[{-Sin[t],Cos[t],t/10},{t,0,a},
    PlotStyle→{Thickness[0.01],Black},ImageSize{250,250}],{{a,10Pi},1.5Pi,10Pi}]
```

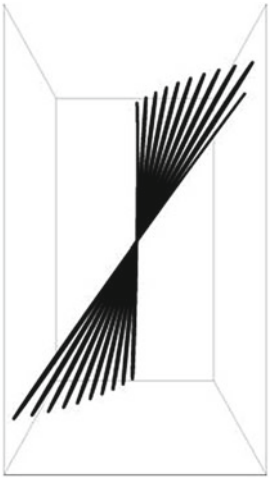


Suggestion: Write a *Mathematica* code to form right- and left-handed helices that are mirror images of each other.

M4. Optical Rotation

This *Mathematica* code simulates a wave approaching the observer and gradually rotating its plane of polarization as it passes through an optically active medium. The wave is built by adding left- and right-circularly polarized waves with slightly different phases:

```
a = 1;      (* phase of left-circularly polarized wave  *)
b = 0.98;   (* phase of right-circularly polarized wave *)
ParametricPlot3D[{Sin[a u]-Sin[b u],Cos[a u]+Cos[b u],(a+b)u/10},
    {u,0,20Pi},Axes→{False,False,False},Boxed→True,PlotStyle→{Thickness[0.01],Black},
    ImageSize→{350,350},ViewPoint→Above]
```



Suggestion: Change the code so that the incoming wave rotates to the counterclockwise direction.

Glossary

Cahn–Ingold–Prelog rules Set of rules named after organic chemists S. Cahn (1899–1981), C.K. Ingold (1893–1970), and V. Prelog (1906–1998) that assign different priorities to the atoms or groups bonded to a chiral center. These priorities follow the decreasing order of the atomic numbers, and the atom or group with lowest priority points away from the observer. If the

	decreasing priority of the atoms or groups bonded to the chiral center and closer to the observer is in the clockwise direction, the chiral center is assigned an *R* descriptor, whereas for the counterclockwise direction, the chiral center is assigned an *S* descriptor
Conformers	Isomers that can be interconverted by internal rotation around a single bond
Electron probability density	Probability of finding any of the n electrons of a molecule in volume element dv, for any position of the remaining electrons and any spin of all electrons
Electrostatic potential	Work done by an electric field to transport a test charge (an infinitesimal positive charge) from infinity to a specified point in a charge distribution divided by the test charge. The electrostatic potential for a molecule is usually mapped over an isodensity surface with a specified isovalue (for example, 0.00040) using rainbow colors, where dark blue is used when the electrostatic potential reaches the extreme positive value (electrophilic region), and dark red indicates the extreme negative value of the electrostatic potential (nucleophilic region)
Enantiomers	Chiral molecules that are mirror images of each other
Intrinsic reaction coordinate	Elementary reaction coordinate that starts at the saddle point and progresses down the hill toward the reactant channel (negative values of the reaction coordinate) and toward the product channel (positive values of the reaction coordinate)
Molecular geometry optimization	Calculation for determining the geometry of a molecule that consists in systematically varying the geometric variables (bond lengths, bond angles, dihedral angles) until an energy minimum is reached. This type of molecular calculation includes two steps, the first leading to a stationary point (a minimum or a maximum), the second step being a frequency calculation to verify whether the stationary point corresponds to a minimum. If first derivatives increase in the vicinity of the stationary point (positive second-order

	derivatives), the stationary point corresponds to an energy minimum, and then the frequency calculation will not produce any imaginary vibrational frequency (in *Gaussian*, imaginary frequencies are presented as negative values)
Mulliken population analysis method	Method developed by Mulliken for apportioning the total electron charge of a molecule to its atoms and overlap regions
Natural bond orbital method	Population analysis method that apportions the electron charge of a molecule to Lewis-like electron pairs localized in chemical bonds and atoms
Natural orbital	Molecular orbital localized in a bond or an atom with maximum electron occupation allowed by the Pauli exclusion principle (number of electrons = 2)
Optical activity	Ability of a chiral molecule to rotate the plane of polarization of light
Polarimeter	Instrument for determining the effect of a substance in rotating the plane of polarization of light. A polarimeter essentially consists of a light source, a polarizer, a sample tube, and a polarization analyzer
Population analysis	Multiple ways of apportioning the total electron charge of a molecule to different regions of space in the molecule, such as atoms and overlap regions (Mulliken population analysis method) or atoms and bonding regions (natural bond orbitals method)
Potential energy surface	A surface that gives the potential energy of a system of three atoms A, B, and C as a function of the internuclear distances A–B and B–C, with the bond angle $\angle ABC$ being kept constant. A typical potential energy surface shows two valleys (the reactant channel and the product channel) approximately parallel to the corresponding Cartesian coordinate axes and an upper pass between them through one first-order saddle point (see Fig. 5.14). The minimum energy path converts reactants into products, passes through the first-order saddle point, and defines the reaction coordinate

Exercises

E1. Explain the expression for the electrostatic potential of a molecule.
E2. Consider a calculation for the H_2 molecule with a minimal basis set and find the coefficients of the 1*s* orbitals in terms of the overlap integral between these orbitals.
E3. Consider the HF/STO-3G calculation on the CH_4 molecule. Starting with the coefficient matrix, use *Mathematica* to obtain the corresponding density matrix.
E4. Consider the density matrix and the full Mulliken population matrix of the *Gaussian* HF/STO-3G calculation on the CH_4 molecule. Find the overlap integral between the $1s_2$ and $1s_3$ hydrogen orbitals.
E5. Consider the full Mulliken population matrix of the *Gaussian* HF/STO-3G calculation on the CH_4 molecule. Find the populations condensed to the carbon atom and to the overlap region between the carbon atom and hydrogen atom 1 ($1s_1$ orbital).
E6. Derive the value for the tetrahedral angle.
E7. Write expressions for the sp^3 hybrid orbitals centered on atom C of CH_4 and show that these hybrids are orthonormal.
E8. Use symmetry and *Mathematica* to write the expressions for the sp^2 hybrid orbitals centered on atom B of BF_3 and show that these hybrids are orthonormal.
E9. Write expressions for the *sp* hybrid orbitals centered on atom C of CO_2 and show that these hybrids are orthonormal.
E10. Derive (5.29).
E11. The analysis of the vibrational spectra of 1,2-dichloroethane shows that the population of the most abundant antiperiplanar conformer is about 80 % at ambient temperature (El Youssoufi et al. 1998). Assuming that only conformers 1 and 2 exist, calculate the energy difference between these conformers at ambient temperature.
E12. The molecular formula for tartaric acid is HOOC–CH(OH)–CH(OH)–COOH. Identify and classify the stereoisomers for this molecule.

References

Mulliken RS (1955) Electronic population analysis on LCAO-MO molecular save functions I. J Chem Phys 23:1833–1840
Baggott JE, Caldow GL, Mills IM (1988) J Chem Soc Faraday Trans 2(84):1407–1422
El Youssoufi Y, Herman M, Liévin J (1998) The ground electronic state of 1,2-dichloroethane I. Ab initio investigation of the geometrical, vibrational and torsional structure. Molec Phys 94:461–472

Further Reading

Levine IN (2013) Quantum chemistry, 7th edn. Pearson Education Inc., Upper Saddle River
Weinhold F, Landis CR (2012) Discovering chemistry with natural bond orbitals. Wiley, Hoboken

Crystals

6

Abstract

We begin this chapter by describing modes for packing disks and spheres, with particular attention to three-dimensional close-packed modes. Next, we consider the concept of translation symmetry, central to crystallography, and present two- and three-dimensional Bravais lattices. We show the distribution of the most representative structures of metals in the periodic table of the elements and build the cesium chloride, sodium chloride, diamond, and zinc blende structures, using their corresponding translation vectors. X-ray diffraction is the most widely used experimental method for solving the structure of crystals. For this reason, we give a brief introduction to the subject and explain the optical transform of an X-ray diffraction experiment. Next, we consider and illustrate the formation of bands in solids, introduce semiconductors and the Fermi–Dirac distribution of electrons for metallic conductors. At the end of this chapter, the student can find several Mathematica codes (Packing of Disks, Hexagon of Disks, Disk Layers, The Third Dimension, HCP Structure, 2D Square Point Lattice, 2D Hexagonal Point Lattice, 3D Bravais Lattices, CsCl Structure, NaCl Structure, From Energy Levels to Bands, Fermi–Dirac Distribution) that enable one to visualize the crystalline structures considered in this chapter and include detailed explanations of new commands, a glossary of important scientific terms, and a list of exercises, whose answers can be found after the Appendix.

6.1 Packing Disks and Spheres

6.1.1 Disks

Consider the packing of identical disks in a plane. Assuming disks of radius 1/2, the distance between centers of tangent disks is 1. In square packing, the (x,y) coordinates

J.J.C. Teixeira-Dias, *Molecular Physical Chemistry*,
DOI 10.1007/978-3-319-41093-7_6

of the disks' centers describe a square (Fig. 6.1). For hexagonal close packing, the evaluation of coordinates of the disks' centers requires simple trigonometry of right angle triangles, as shown in Fig. 6.1.

Figure 6.2 shows square and hexagonal packing modes of identical disks. The *Mathematica* code **M1** builds the disk lattices of this figure. In square packing, each disk is tangent to four neighboring disks, whereas in hexagonal packing, each disk is tangent to six neighboring disks (Fig. 6.2) whose centers are the vertices of a regular hexagon. This arrangement reveals the existence of a sixfold rotational symmetry axis. The *Mathematica* code **M2** illustrates this point and generates the picture shown in Fig. 6.3.

As shown in Fig. 6.4, the centers of three close-packed disks are the vertices of an equilateral triangle, whose center is the center of the interstice that can be determined making use of simple trigonometry for a right triangle.

The *Mathematica* code **M3** considers two close-packed layers of identical disks, where the second layer is displaced so that its disks' centers coincide with the

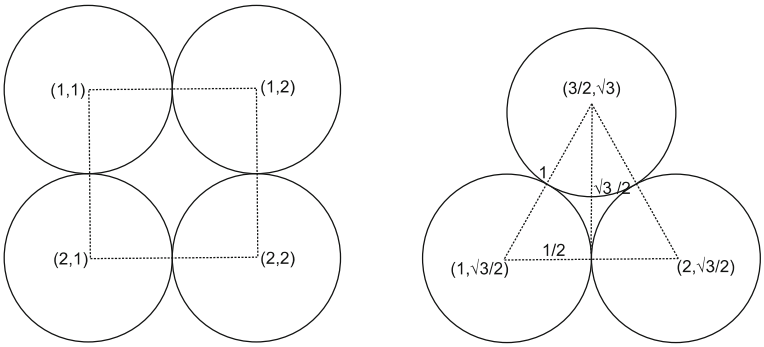


Fig. 6.1 Coordinates of centers of disks of radius 1/2, in square and hexagonal packing modes. Figure obtained with *Mathematica*

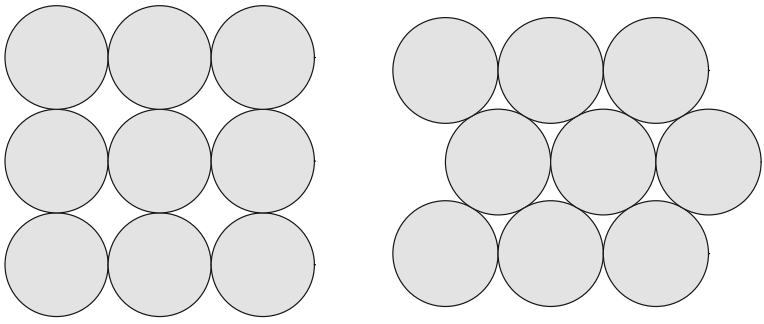


Fig. 6.2 Square and hexagonal packing modes of identical disks. Figure obtained with *Mathematica*

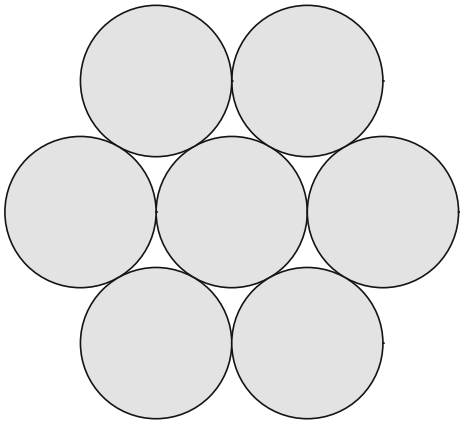


Fig. 6.3 Disks exhibiting sixfold rotational axis for close-packed identical disks on the plane. Figure obtained with *Mathematica*

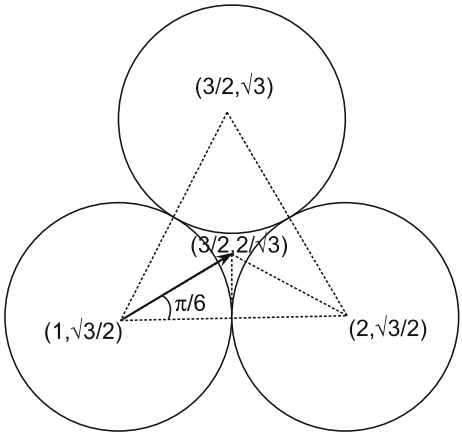


Fig. 6.4 (x, y) coordinates for the center of the interstice of three close-packed identical disks in a plane. Disks drawn with *Mathematica*

interstices of the first layer (Fig. 6.5). The translation vector that shifts the second layer of disks with respect to the first layer is shown in Fig. 6.4. The coordinates of its origin are $(1, \sqrt{3}/2)$, and the coordinates of its tip at the center of the interstice are $(3/2, 2/\sqrt{3})$. Hence, the translation vector is given by

$$\mathbf{t} = \frac{1}{2}\hat{\mathbf{x}} + \frac{\sqrt{3}}{6}\hat{\mathbf{y}} \tag{6.1}$$

6.1.2 Spheres

We begin by considering four equivalent and mutually tangent spheres of radius 1/2 whose centers form a regular tetrahedron that can be drawn inside a cube (Fig. 6.6). The distance between center 4 and the 1-2-3 plane is given by $\sqrt{2}/\sqrt{3}$, as can be

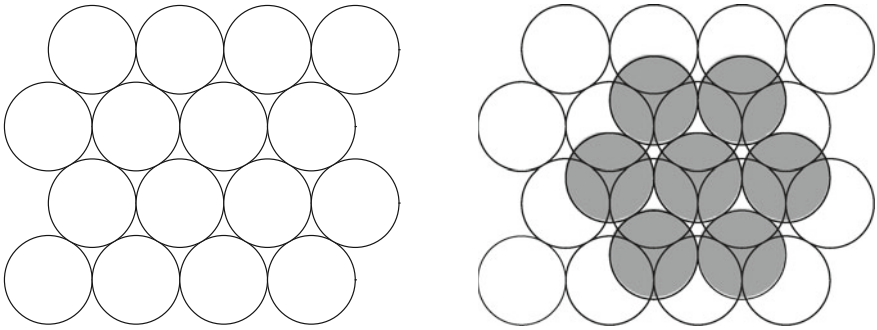


Fig. 6.5 Close-packed disk layer and a second close-packed layer of disks in *gray* displaced so that its disks' centers coincide with the interstices of the first layer. Figure obtained with *Mathematica*

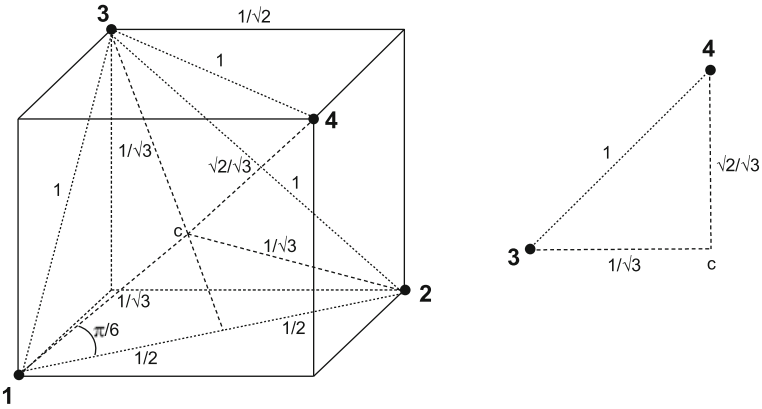


Fig. 6.6 Right triangle trigonometry determines the distance between center 4 and the 1-2-3 plane

shown by simple trigonometry of right triangles (Fig. 6.6). Sphere 4 is nested on the interstice of spheres 1, 2, and 3. Therefore, $\sqrt{2}/\sqrt{3}$ is the vertical distance between the centers of two horizontal close-packed layers of spheres of radius 1/2. The vector that goes from center 1 to center 4 [= vector (4-c) + vector (c-1), where vector (4-c) is obtained from Fig. 6.6 and vector (c-1) is given by (6.1)] is

$$\mathbf{t} = \frac{1}{2}\hat{\mathbf{x}} + \frac{\sqrt{3}}{6}\hat{\mathbf{y}} + \sqrt{\frac{2}{3}}\hat{\mathbf{z}} \tag{6.2}$$

(**E1**).

Figure 6.7 shows a regular tetrahedron formed by the centers of four mutually tangent spheres. The *Mathematica* code **M4** uses the threefold rotation symmetry axis to draw Fig. 6.7 (**E2**).

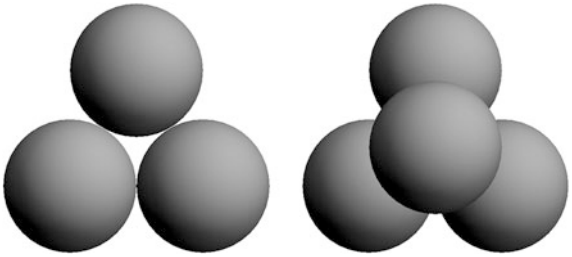


Fig. 6.7 Regular tetrahedron formed by the centers of four mutually tangent spheres. Figure obtained with *Mathematica*

6.1.3 Hexagonal Close Packing

The close packing of identical spheres on one layer poses no great difficulty: each sphere is tangent to six neighboring spheres, whose centers form a regular hexagon. If projected on the plane of the paper, the spheres are seen as tangent disks with interstices appearing as triangles of curved sides, half of them with an upside vertex, the other half with a downside vertex (Fig. 6.8). The second layer of spheres nests on interstices with downside vertices, since these correspond to depressions on the upper surface of the layer of spheres. Unlike the first layer, where the downside and upside curved triangles are equivalent sites (rotation of the whole structure by 180° transforms downside curved triangles into upside curved triangles and vice versa), the second layer has two kinds of interstices: those that pass through both layers and those that encounter spheres of the first layer (Fig. 6.8).

Given the option to lay down the third layer, we nest the spheres of this layer on the second layer interstices that encounter first layer spheres, keeping open the interstices that pass through all layers. Then the third and first layers coincide on projection (Fig. 6.8). If successive layers are labeled by capital letters with the first layer being labeled A, then the second layer is B and the third layer is A again (layer sequence ABAB… with spheres of alternating layers overlaying one another). This structure is called **hexagonal close-packed** (**hcp**), with the representative polyhedron being a hexagonal prism (Fig. 6.8). The *Mathematica* code **M5** builds Fig. 6.8.

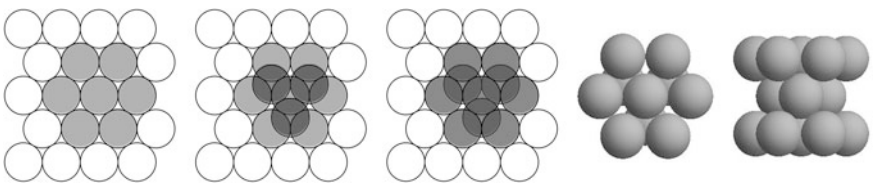


Fig. 6.8 Hexagonal close packing of rigid spheres with equal radius. Figure obtained with *Mathematica*

The number of spheres tangent to any one sphere (number of nearest neighbors) is called **coordination number**. The coordination number of the hcp structure is 12. To obtain this number, we consider layer A and two additional B layers, one above, the other below, to form the BAB sequence. The central sphere of layer A is tangent to six spheres on the same equatorial layer and to three spheres on each of the B layers, above and below the central sphere of layer A.

6.1.4 Cubic Close Packing

We now go back to the first two layers of close-packed spheres and nest the sphere of the third layer on the interstice that passes through B and A (Fig. 6.9). The resulting sequence of layers ABCABC… leads to the **cubic close-packed** (**ccp**) structure, which is also a **face-centered cubic** (**fcc**) structure, since the representative polyhedron is a face-centered cube. Layers A and B correspond to planes that pass through face diagonals of three cube faces. The central sphere of the cube top face is tangent to four spheres on this face, four spheres on the layer below, and four spheres on the layer above (the latter is not shown in Fig. 6.9) thus having the coordination number 12.

6.1.5 Packing Densities

Spheres cannot occupy 100 % of the volume of a crystal repeating unit. The efficiency of occupation of the three-dimensional space by rigid spheres can be measured by the **packing density**, which is a ratio of two volumes, the volume of all spheres contained in the repeating unit (a parallelepiped; note that the cube is a particular kind of parallelepiped) over the volume of the repeating unit. In order to evaluate the packing density, one should take into account the contribution of each sphere to the repeating unit of the crystal, which depends on the position of the sphere in the parallelepiped of the repeating unit. Each sphere located on a parallelepiped vertex is shared by eight repeating units (Fig. 6.10). Thus, each repeating unit contains 1/8 of one sphere. If the sphere occupies the midpoint of one edge, then each parallelepiped contains 1/4 of the sphere. In turn, if the sphere occupies the center of a face, it is shared by two parallelepipeds, so 1/2 belongs to each parallelepiped.

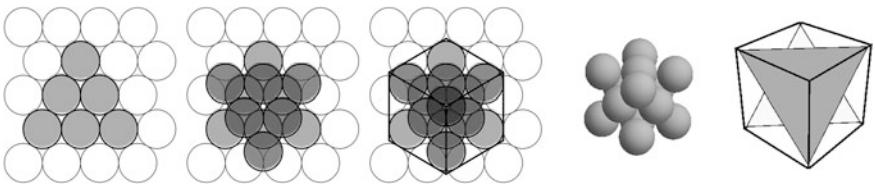


Fig. 6.9 Cubic close packing of spheres with equal radius. Figure obtained with *Mathematica*

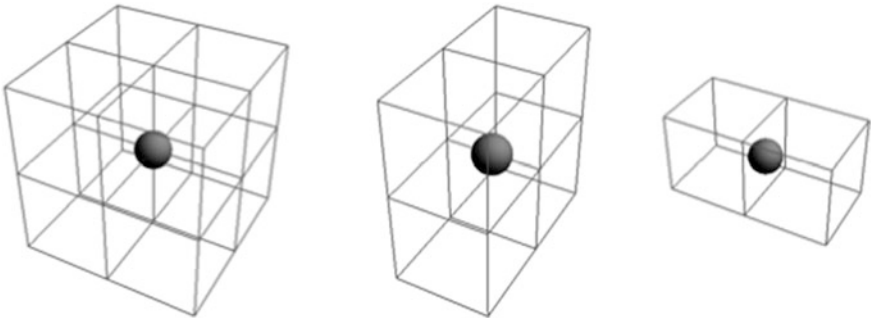


Fig. 6.10 Spheres occupying a vertex, the midpoint of an edge, and the center of a face contribute to each crystal repeating unit with 1/8, 1/4, and 1/2 of a sphere, respectively

If one layer of the hcp structure (say layer A) is displaced to nest over yet uncovered interstices of the preceding layer, then it transforms itself into layer C of the ccp structure, and the ABA sequence of the original hcp structure is converted into the ABC sequence of the ccp structure. This reasoning suggests that the hcp and ccp structures have the same coordination number (6.12) and the same packing density, since they are both close-packed structures differing by displaced layers of spheres.

In 1611, the astronomer and mathematician Johannes Kepler (1571–1630) conjectured that it would be impossible to exceed the packing density of the cubic close-packed structure. In the nineteenth century, Karl Friedrich Gauss (1777–1855) showed that the hexagonal close-packed structure reached the maximum packing density for a lattice of identical spheres. Finally, in 1998, Thomas Hales proved Kepler's conjecture, thus concluding that both the hexagonal and the cubic close-packed structures reach the maximum packing density for a regular structure of identical spheres.

In the **cubic close-packed** (**ccp**) structure, the central sphere of a cube face is tangent to the spheres on the vertices (Fig. 6.11). Therefore, we can write

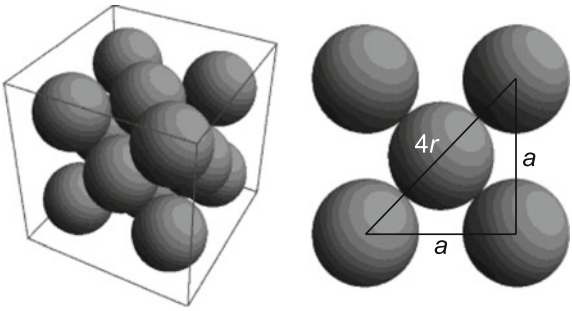


Fig. 6.11 In the cubic close-packed structure, spheres are tangent along the face diagonals. Arrangement of spheres obtained with *Mathematica*

$$4r = \sqrt{a^2 + a^2} \quad \therefore \quad a = 2\sqrt{2}\, r \tag{6.3}$$

where a is the edge and r is the sphere's radius. The packing density is given by

$$\frac{\left(8 \times \frac{1}{8} + 6 \times \frac{1}{2}\right) \frac{4}{3} \pi r^3}{a^3} = \frac{4 \times \frac{4}{3} \pi r^3}{\left(2\sqrt{2}\, r\right)^3} = \frac{\pi}{3\sqrt{2}} \approx 0.74 \tag{6.4}$$

(**E3**).

The **body-centered cubic** (**bcc**) structure has a coordination number equal to 8, since the central sphere is tangent to eight spheres on the vertices of the cube (Fig. 6.12). The right triangle formed by one cube main diagonal, one cube face diagonal, and one cube edge allows one to determine the cube edge a as a function of the sphere's radius r:

$$4r = \sqrt{a^2 + d_{\text{face}}^2} = \sqrt{a^2 + 2\,a^2} = \sqrt{3}\, a \quad \therefore \quad a = \frac{4}{\sqrt{3}} r \tag{6.5}$$

Thus, the packing density is given by

$$\frac{\left(8 \times \frac{1}{8} + 1\right) \frac{4}{3} \pi r^3}{a^3} = \frac{2 \times \frac{4}{3} \pi r^3}{\left(\frac{4}{\sqrt{3}} r\right)^3} = \frac{\sqrt{3}\, \pi}{8} \approx 0.68 \tag{6.6}$$

This packing density is smaller than those of the hcp and ccp structures: the bcc structure is not a close-packed structure.

The **simple cubic** (**sc**) structure presents spheres on the cube vertices (Fig. 6.13), and each sphere is tangent to six spheres on the vertices of a regular octahedron. The packing density is given by

$$\frac{\left(8 \times \frac{1}{8}\right) \frac{4}{3} \pi r^3}{a^3} = \frac{\frac{4}{3} \pi r^3}{(2r)^3} = \frac{\pi}{6} \approx 0.52 \tag{6.7}$$

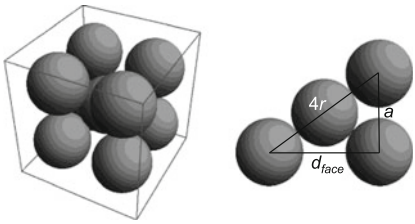


Fig. 6.12 In the body-centered cubic structure, spheres are tangent along the cube's main diagonals. Arrangement of spheres obtained with *Mathematica*

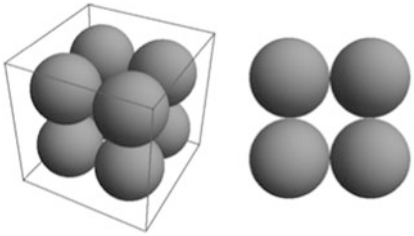


Fig. 6.13 Spheres are tangent along the cube edges of the simple cubic structure. Figure obtained with *Mathematica*

The coordination numbers and packing densities of the above-mentioned structures of identical spheres are summarized in the following table:

Structure	Coordination number	Packing density
Hexagonal close-packed (hcp)	12	0.74
Cubic close-packed (ccp)	12	0.74
Body-centered cubic (bcc)	8	0.68
Simple cubic (sc)	6	0.52

6.1.6 Occupying Interstices

The interstices of close-packed structures of identical spheres are of two types: octahedral, occupying the center of a regular octahedron whose vertices are the centers of six spheres, and tetrahedral, occupying the center of a regular tetrahedron whose vertices are the centers of four spheres (Fig. 6.14).

Imagine the **octahedral interstice** occupied by a small sphere tangent to the surrounding larger spheres. The ratio of the smaller radius and the larger radius can be easily obtained from Fig. 6.15 and is given by

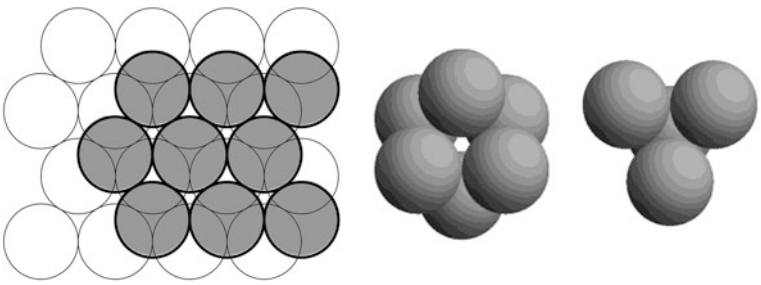


Fig. 6.14 Close-packed layers of identical spheres have octahedral and tetrahedral interstices. Figure obtained with *Mathematica*

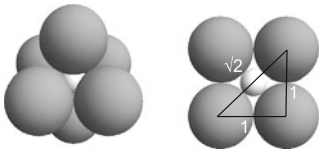


Fig. 6.15 Octahedral interstice occupied by one sphere. The diameter of each larger sphere is arbitrarily equal to 1. Figure obtained with *Mathematica*

$$\frac{r_s}{r} = \frac{2r_s}{2r} = \frac{\sqrt{2}-1}{1} \approx 0.414 \tag{6.8}$$

where r_s is the radius of the smaller sphere that occupies the octahedral interstice.

Consider now the **tetrahedral interstice** occupied by a sphere tangent to its neighboring spheres. From the right triangle of Fig. 6.16, we can write

$$1 = (1 + r_s)\sin(\alpha/2) \quad \therefore \quad r_s = \frac{1}{\sin(\alpha/2)} - 1 \tag{6.9}$$

and consequently,

$$\frac{r_s}{r} = \frac{\frac{1}{\sin(\alpha/2)} - 1}{1} = \frac{1}{\sin\left[\arccos(-1/3)/2\right]} - 1 \approx 0.225 \tag{6.10}$$

where α is the tetrahedral angle, and the radii of the larger spheres are set arbitrarily equal to 1.

Let us now consider the simple cubic structure and the occupation of the **cubic interstice** by a sphere tangent to its eight neighboring spheres. The ratio of the radius of the smaller sphere and the radius of the larger sphere can be easily obtained from Fig. 6.17 and is given by

$$\frac{r_s}{r} = \frac{2r_s}{2r} = \frac{\sqrt{3}-1}{1} \approx 0.732 \tag{6.11}$$

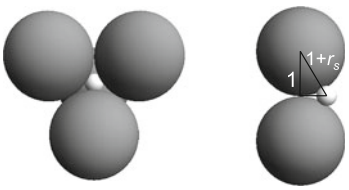


Fig. 6.16 Tetrahedral interstice occupied by one sphere. The *marked angle* is half of the tetrahedral angle. The radii of the larger spheres are set arbitrarily equal to 1. Arrangement of spheres obtained with *Mathematica*

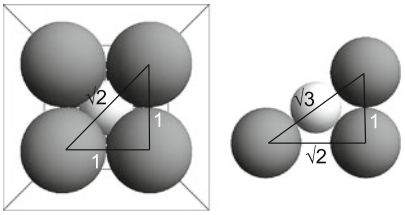


Fig. 6.17 Simple cubic interstice occupied by one sphere. The diameter of each larger sphere is arbitrarily equal to 1. Arrangement of spheres obtained with *Mathematica*

The values of the ratios r_s/r and the corresponding coordination numbers are collected in the following table:

r_s/r	Environment	Coordination number
0.225 − (0.414)	Tetrahedral	4
0.414 − (0.732)	Octahedral	6
0.732 − 1.000	Cubic	8

The lower value of each r_s/r range has been determined assuming that the spheres of larger radii preserve their initial structure and the sphere of smaller radius is tangent to its surrounding spheres. In each r_s/r range, values greater than the lower value force the larger spheres to drift away from each other, no longer being tangent. At each upper limit in parentheses, the structure abruptly changes, and its coordination number increases.

6.2 Translation Symmetries

In this section, we consider the concept of translation symmetry, central to crystallography. We begin with a discrete two-dimensional array of equivalent points that describes the periodicity of the original crystal structure, regardless of the physical content (spheres, atoms, molecules) ascribed to its points, since these are devoid of any physical content and simply emphasize the translation symmetry.

6.2.1 2D Bravais Lattices

Figure 6.18 shows a finite section of an infinite discrete array of points with a square arrangement. For each point P of this array, it is possible to find a pair of integers (n_1, n_2) (positive or negative integers and zero included) such that the positional vector of a point P is given by

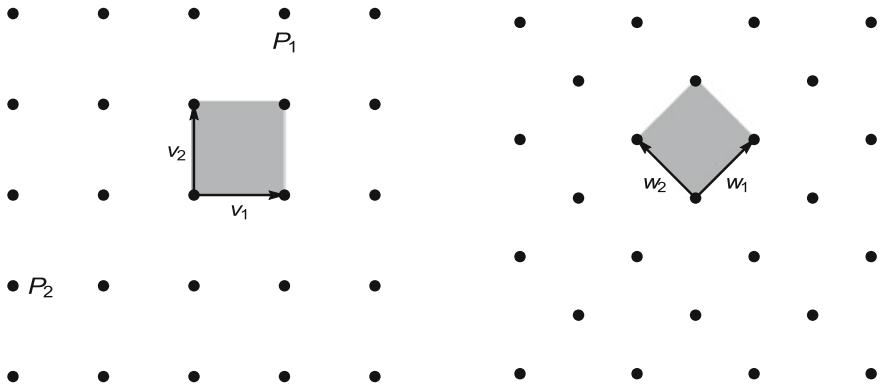


Fig. 6.18 Square point lattice and its π/4 counterclockwise rotation. Figure obtained with *Mathematica*

$$\mathbf{p} = n_1\mathbf{v}_1 + n_2\mathbf{v}_2 \tag{6.12}$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent vectors (noncollinear vectors). Each point on the lattice corresponds to a defined pair of integers (n_1, n_2). For instance, $P_1 = \mathbf{v}_1 + 2\mathbf{v}_2$ and $P_2 = -2\mathbf{v}_1 - \mathbf{v}_2$ correspond to (n_1, n_2) equal to (1, 2) and (−2, −1), respectively. The vector Eq. (6.12) generates every point of the 2D square array. Therefore, all points are equivalent, since every point has the same arrangement of neighbors. In the case of Fig. 6.18, $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal and are given by

$$\mathbf{v}_1 = a\,\hat{\mathbf{x}} \quad \mathbf{v}_2 = a\,\hat{\mathbf{y}} \tag{6.13}$$

where the scalar a is the **lattice constant**, vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are called **primitive translation vectors**, and the array of points is a two-dimensional **Bravais square lattice**. All points that satisfy the following equality cover the gray area:

$$\mathbf{r} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 \quad 0 \le x_i \le 1 \tag{6.14}$$

When this area is translated by all vectors **p**, the two-dimensional space is filled without overlapping itself or leaving voids, and such an area is called a two-dimensional **unit cell**. The *Mathematica* code **M6** uses (6.12) and (6.13) to build a square point lattice like the one of Fig. 6.18 and applies a π/4 counterclockwise rotation to the resulting square point lattice.

Both arrays of points of Fig. 6.18 represent the same lattice, since the translation vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ result from $\mathbf{v}_1$ and $\mathbf{v}_2$ by a π/4 counterclockwise rotation,

$$\mathbf{w}_1 = R_{\pi/4}\mathbf{v}_1 \quad \mathbf{w}_2 = R_{\pi/4}\mathbf{v}_2 \tag{6.15}$$

where $R_{\pi/4}$ is the rotation matrix

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{6.16}$$

for $\theta = \pi/4$,

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6.17}$$

and

$$\mathbf{w}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \mathbf{w}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \tag{6.18}$$

From (6.17) and (6.18), we obtain

$$\mathbf{w}_1 = \frac{1}{\sqrt{2}}(\mathbf{v}_1 + \mathbf{v}_2) \quad \mathbf{w}_2 = \frac{1}{\sqrt{2}}(\mathbf{v}_2 - \mathbf{v}_1) \tag{6.19}$$

where $1/\sqrt{2}$ is the normalizing factor ($\mathbf{v}_1$ and $\mathbf{v}_2$ are normalized vectors).

The choice of primitive translation vectors is not unique. For instance, we can take $\mathbf{v}_1$ and $\mathbf{v}_1 + \mathbf{v}_2$ (Fig. 6.19), since these are linearly independent vectors. However, $\mathbf{v}_1$ and $\mathbf{v}_1 + \mathbf{v}_2$ are not orthogonal, do not have the same length, and so do not reflect the symmetry of the square lattice. Therefore, they are not a convenient choice of primitive vectors for a square array.

Consider now a two-dimensional lattice where each point is surrounded by six points on the vertices of a regular hexagon (Fig. 6.20). The translation vectors of this **hexagonal** lattice are given by

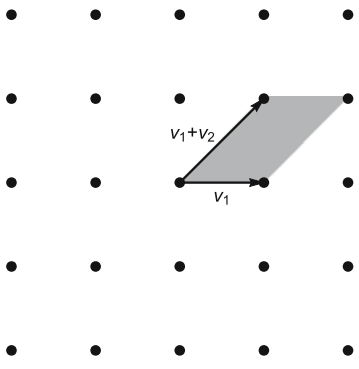


Fig. 6.19 The translation vectors $\mathbf{v}_1$, $\mathbf{v}_1 + \mathbf{v}_2$ do not reflect the symmetry of a simple square lattice. Figure obtained with *Mathematica*

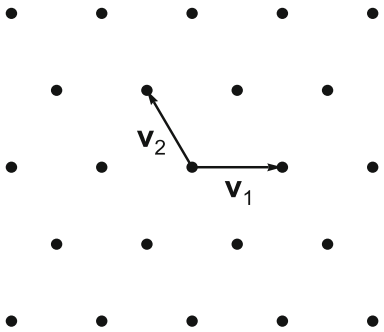


Fig. 6.20 Two-dimensional hexagonal Bravais lattice. Figure obtained with *Mathematica*

$$\mathbf{v}_1 = a\,\hat{\mathbf{x}} \quad \mathbf{v}_2 = a\left(-\frac{1}{2}\hat{\mathbf{x}} + \frac{\sqrt{3}}{2}\hat{\mathbf{y}}\right) \tag{6.20}$$

where a is the lattice constant and these vectors correspond to the sides of a unit-cell parallelogram. The *Mathematica* code **M7** uses (6.12) and (6.20) to build a hexagonal point lattice like the one of Fig. 6.20 (**E4**).

In two-dimensional space, there are five Bravais lattices, namely square, rectangular, oblique, hexagonal, and centered rectangular lattices (Fig. 6.21).

The black and gray points of Fig. 6.22 correspond to motifs of a distinct nature, where each black point has four gray points as closest neighbors, and conversely, each gray point has four black points as closest neighbors. This structure cannot be described by a single set of primitive translation vectors, since the points marked black and gray cannot simultaneously correspond to integer values of n_1 and n_2 in (6.12). The dark gray vector is given by

$$\mathbf{t} = \frac{\mathbf{v}_1 + \mathbf{v}_2}{2} \tag{6.21}$$

When this vector is applied to every point of the black lattice, it generates the gray lattice and vice versa. One says that the periodic structure of Fig. 6.22 results from merging two square point lattices. The resulting structure can be considered a two-dimensional equivalent of the cesium chloride structure. The tri-dimensional

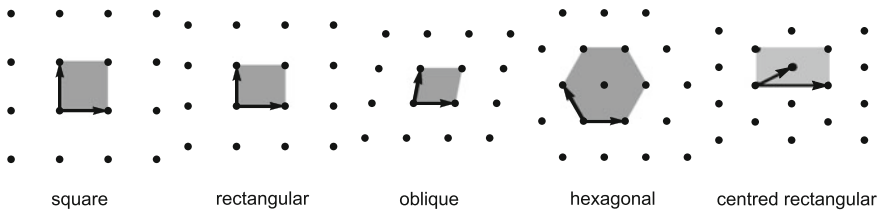


Fig. 6.21 Bravais lattices in two-dimensional space. Lattices and translation vectors obtained with *Mathematica*

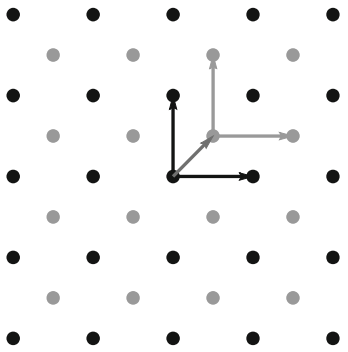


Fig. 6.22 Two square lattices merge with one another. When the orange translation vector is applied to every point on the *black* point lattice, the *gray* lattice is obtained. Figure obtained with *Mathematica*

CsCl structure results from merging two simple cubic lattices, one for cesium ions, the other for chloride ions (cesium and chloride ions have approximately the same radius).

6.2.2 3D Bravais Lattices

A three-dimensional Bravais lattice is an infinite array of discrete points such that for every point P of the array defined by the positional vector $\mathbf{p}$, it is possible to find a set of positive or negative integers (zero included) (n_1, n_2, n_3) that satisfy the following vector equation:

$$\mathbf{p} = n_1\mathbf{v}_1 + n_2\mathbf{v}_2 + n_3\mathbf{v}_3 \tag{6.22}$$

where $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are linearly independent translation vectors (noncollinear vectors) that span the lattice. The lengths of cell edges (a, b, c) and the angles between them (α, β, γ) characterize the lattice systems.

In three-dimensional space, there are 7 lattice systems and 14 Bravais lattices. The lattice systems are cubic (a, a, a; $\alpha = \beta = \gamma = 90°$), tetragonal ($a$, a, c; $\alpha = \beta = \gamma = 90°$), orthorhombic ($a$, b, c; $\alpha = \beta = \gamma = 90°$), monoclinic ($a$, b, c; $\alpha = \gamma = 90°$, $\beta \neq 90°$), triclinic (a, b, c; α, β, $\gamma \neq 90°$), trigonal (a, a, a; α, β, $\gamma \neq 90°$), and hexagonal (a, a, c; $\alpha = \beta = 90°$, $\gamma = 120°$). The Bravais lattices are simple cubic, body-centered cubic, face-centered cubic, simple tetragonal, body-centered tetragonal, simple orthorhombic, base-centered orthorhombic, body-centered orthorhombic, face-centered orthorhombic, simple monoclinic, base-centered monoclinic, simple triclinic, simple trigonal, and simple hexagonal (see *Mathematica* code **M8**). To extend group theory considerations from the molecule (point groups; see Chap. 4) to the crystalline state (space groups), translations need to be considered.

We now illustrate the translation vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ for a few crystal structures. Starting with the **simple cubic** structure, its translation vectors are given by

$$\mathbf{v}_1 = a\,\hat{\mathbf{x}} \quad \mathbf{v}_2 = a\,\hat{\mathbf{y}} \quad \mathbf{v}_3 = a\,\hat{\mathbf{z}} \tag{6.23}$$

as suggested by the symmetry of the structure (Fig. 6.23, **E5**). Polonium provides a rare example of the simple cubic structure.

We now consider the **body centered cubic** structure, whose translation vectors are given by

$$\mathbf{v}_1 = a\,\hat{\mathbf{x}} \quad \mathbf{v}_2 = a\,\hat{\mathbf{y}} \quad \mathbf{v}_3 = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}) \tag{6.24}$$

where $\mathbf{v}_3$ is along the cube main diagonal (Fig. 6.24, **E6**), and the projection of $\mathbf{v}_3$ along, $\hat{\mathbf{x}}, \hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ has length $a/2$. The alkali metal cesium is an example of the body-centered cubic structure.

The translation vectors for the **face-centered cubic** structure are given by

$$\mathbf{v}_1 = \frac{a}{2}(\hat{\mathbf{y}} + \hat{\mathbf{z}}) \quad \mathbf{v}_2 = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{z}}) \quad \mathbf{v}_3 = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}}) \tag{6.25}$$

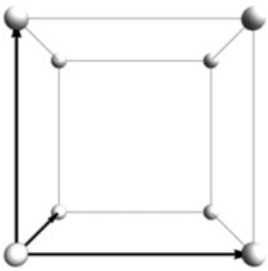


Fig. 6.23 Translation vectors for the simple cubic structure. Figure obtained with *Mathematica*

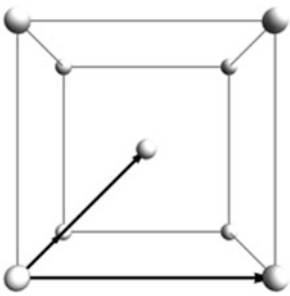


Fig. 6.24 Translation vectors for the body-centered cubic structure. Figure obtained with *Mathematica*

(Figure 6.25, **E7**). Aluminum gives an example of the face-centered cubic structure.

The **hexagonal close-packed** structure is formed by two Bravais lattices, one spanned by the translation vectors

$$\mathbf{v}_1 = a\,\hat{\mathbf{x}} \quad \mathbf{v}_2 = a\left(\frac{1}{2}\hat{\mathbf{x}} + \frac{\sqrt{3}}{2}\hat{\mathbf{y}}\right) \quad \mathbf{v}_3 = c\,\hat{\mathbf{z}} \tag{6.26}$$

the other obtained by translation by the vector

$$\mathbf{t} = \frac{\mathbf{v}_1}{3} + \frac{\mathbf{v}_2}{3} + \frac{\mathbf{v}_3}{2} \tag{6.27}$$

(Figure 6.26, **E8**). Zinc is an example of the hexagonal close-packed structure.

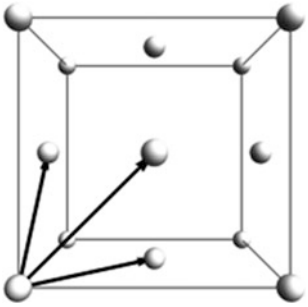


Fig. 6.25 Translation vectors for the face-centered cubic structure. Figure obtained with *Mathematica*

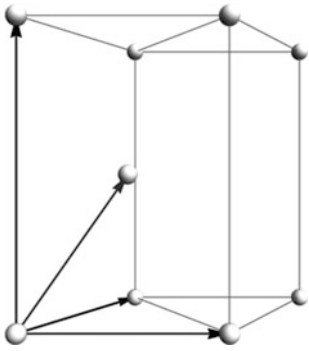


Fig. 6.26 Translation vectors for the hexagonal close-packed structure. Figure obtained with *Mathematica*

6.3 Crystal Structures

6.3.1 Metals

The hexagonal close-packed (hcp), cubic close-packed (ccp) and body-centered cubic (bcc) structures are the most representative structures for the metals of periods 2–6 and groups 1–14 of the periodic table (Fig. 6.27). Mercury, gallium, indium, and tin are not among the metals with these structures. On the whole, there is a tendency for metals of the same group to share the same crystalline structure, as illustrated, for the bcc structure by groups 1 (IA), 5 (VB), and 6 (VIB), for the hcp structure by groups 3 (IIIB), 4 (IVB), and 12 (IIB), and for the ccp structure by groups 10 (VIIIB) and 11 (IB).

6.3.2 Lattice Energy

Consider now an ionic crystal of general formula X_mY_n. The energy of formation of one mole of X_mY_n from infinitely separated ions,

$$mX^{n+}(g) + nY^{m-}(g) \rightarrow X_mY_n(s) \tag{6.28}$$

is called the **lattice energy** and is invariably negative (energy released by the system). The lattice energy (electrostatic energy of attraction and repulsion between ions in one mole of X_mY_n) plus the energy term associated with the volume change at constant pressure (the ions X^{n+} and Y^{m-} are in the gaseous state) is called the **lattice enthalpy**. This energy is usually determined by the Born–Haber cycle, a multistep thermochemical cycle in which the enthalpy of one of the steps is the lattice enthalpy due to German scientists, Max Born (1882–1950; Nobel Prize in physics in 1954 for his research achievements on the statistical interpretation of quantum mechanics) and Fritz Haber (1868–1934; Nobel prize in chemistry in 1918, for his research on the synthesis of ammonia). When the lattice energy is not

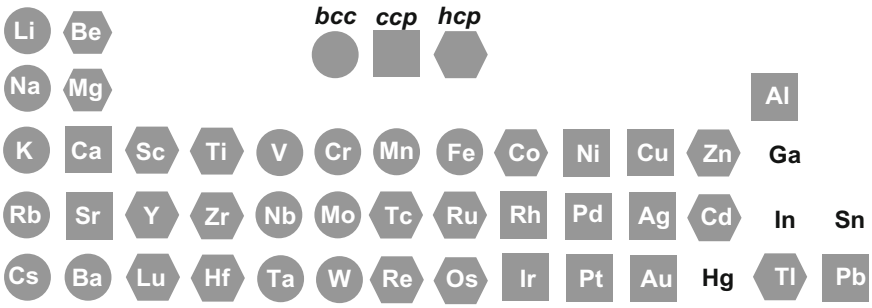


Fig. 6.27 Distribution of body-centered cubic, cubic close-packed, and hexagonal close-packed structures on the periodic table of elements

tabulated, the Kapustinskii equation can be used to obtain a calculated value (Kapustinskii 1956).

6.3.3 Cesium Chloride and Sodium Chloride

Cesium chloride is formed by equal amounts of cesium and chloride ions. Its crystal structure results from merging two simple cubic lattices, one for the Cs^+ ions, the other for the Cl^- ions. Its lattice energy is about -670 kJ mol^{-1} (Handbook of Chemistry and Physics, 2011). If one Cl^- is at the origin, the translation vectors for the chloride ion lattice are given by (6.23), and the vector that translates this lattice to produce the cesium ion lattice is given by

$$\mathbf{t} = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}) \tag{6.29}$$

whose magnitude is equal to half of the main diagonal of a cube with chloride ions at the vertices. The *Mathematica* code **M9** builds the cesium chloride structure with the Cs^+ and Cl^- ions being represented by spheres (Fig. 6.28).

Sodium chloride is formed by equal amounts of sodium and chloride ions. Its crystal structure results from merging two face-centered cubic lattices, one for Na^+ ions, the other for Cl^- ions (Fig. 6.29). Its lattice energy is about -767 kJ mol^{-1} (Handbook of Chemistry and Physics, 2011). If one Cl^- is at the origin, the translation vectors for the chloride ion lattice are given by (6.25), and the vector that translates this lattice to generate the sodium lattice is given by

$$\mathbf{t} = \frac{a}{2}\hat{\mathbf{z}} \tag{6.30}$$

The *Mathematica* code **M10** builds the sodium chloride structure, with the Na^+ and Cl^- ions represented by spheres (Fig. 6.29).

Fig. 6.28 The cesium chloride structure with the Cs^+ and Cl^- ions represented by spheres (Cs^+ *light gray spheres*, Cl^- *gray spheres*). Figure obtained with *Mathematica*

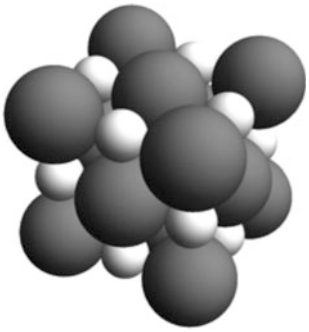


Fig. 6.29 The sodium chloride structure with the Na^+ and Cl^- ions represented by spheres (Na^+ *light gray spheres*, Cl^- *gray spheres*). Figure obtained with *Mathematica*

6.3.4 Diamond and Zinc Blende

In diamond, each carbon atom is covalently bonded to four carbon atoms on the vertices of regular tetrahedra. The diamond lattice consists of two face-centered cubic lattices that interpenetrate each other. The translation vectors for one of these lattices is given by

$$n_1\mathbf{v}_1 + n_2\mathbf{v}_2 + n_3\mathbf{v}_3 \tag{6.31}$$

where $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are given by (6.25), and (n_1, n_2, n_3) are integers (which can be positive, negative, or zero). The translation vectors for the other lattice of diamond is given by

$$n_1\mathbf{v}_1 + n_2\mathbf{v}_2 + n_3\mathbf{v}_3 + \mathbf{t} \tag{6.32}$$

where $\mathbf{t}$ is given by

$$\mathbf{t} = \frac{a}{4}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}) \tag{6.33}$$

The magnitude of $\mathbf{t}$ is one-fourth that of the cube's main diagonal. For the sake of clarity, the spheres that schematically represent carbon atoms in Fig. 6.30 have a radius much smaller than the covalent radius of the carbon atom, and darker and lighter spheres represent carbon atoms belonging to distinct lattices.

Zinc blende and sphalerite are names for a mineral whose composition is mostly zinc sulfide with a variable percentage of zinc being replaced by iron (II). The iron presence gives a dark and glossy appearance to zinc blende, similar to that of galena, an iron (II) sulfide. The likely confusion of zinc blende with galena is at the origin of the word *sphalerite*, from the Greek word *sphaleron*, meaning *misleading*. The word *blende* gives the same idea, but the origin of *blende* is German (*blenden* = *to blind*, also *to deceive*). In this text, *zinc blende* is used to refer to one of the

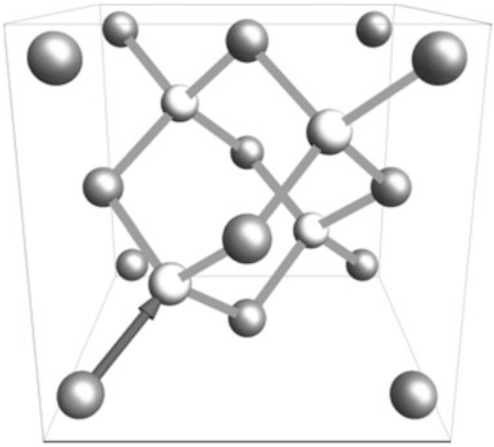


Fig. 6.30 Diamond structure with an illustration of the vector **t**. Figure obtained with *Mathematica*

crystalline forms of zinc sulfide, ZnS, consisting of two merged face-centered cubic lattices, one for the sulfur atoms, the other for the zinc atoms.

Figure 6.31 shows two distinct main cubes that schematically represent the *same* zinc blende crystal. The cube on the right results from the one on the left by shifting the latter by the vector **t**. This vector is directed along the diagonal of the cube, and its length is one-fourth that of the main diagonal [see (6.33) and Fig. 6.31]. The translation of the atoms of the main cube on the left by the vector **t** leads to the zinc blende structure on the right of Fig. 6.31. One of these main cubes (the one on the left) shows sulfur atoms represented by light gray spheres on the vertices and face centers and four zinc atoms (gray spheres) totally contained in the repeating unit. In turn, the cube on the right shows zinc atoms on the vertices and face centers and four sulfur atoms totally contained in the repeating unit.

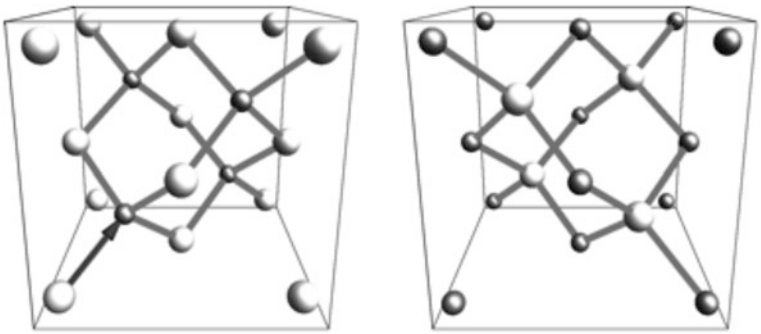


Fig. 6.31 Zinc blende structure with an illustration of the vector **t** in the left-hand drawing: *light-gray spheres* represent sulfur atoms, and *gray spheres* represent zinc atoms. Left- and right-hand drawings represent two main cubes of the same crystal. Figure obtained with *Mathematica*

If we divide each main cube into eight smaller and equal cubes, we conclude that only four of these are occupied by ZnS_4 tetrahedra on the left of Fig. 6.31, and by SZn_4 tetrahedra on the right of this figure. On the whole, each of the main cubes corresponds to the formula Zn_4S_4, since there are $8 \times (1/8) + 6 \times (1/2) = 4$ atoms of one of the elements (sulfur or zinc) on the vertices and centers of the faces and four atoms of the other element totally (zinc or sulfur) included in the repeating unit. If carbon atoms substitute zinc and sulfur atoms and the lattice is correctly adjusted to make all bond lengths equal, one obtains the diamond structure.

6.4 X-Ray Diffraction

Atoms and monatomic ions have dimensions of order 1 Å (= 10^{-10} m), and electromagnetic radiation with wavelength of this order of magnitude consists of X-rays. In 1912, Max von Laue (1879–1960, Nobel Prize in physics 1914) showed that crystals diffract X-rays and at the same time, William Henry Bragg and William Lawrence Bragg, father and son, solved several crystalline structures using X-ray diffraction and shared the Nobel Prize in physics in 1915. One of the first structures to be determined was that of sodium chloride. Nowadays, X-ray diffraction is the most widely used method for solving the structure of crystals, in particular that of macromolecules of biological importance such as proteins.

After being irradiated with X-ray waves, each atom acts like a secondary source of X-rays. When an X-ray beam with defined wavelength impinges on a crystal, multiple X-ray beams with the same wavelength emerge that can leave their imprint (**diffraction pattern**) on an X-ray sensitive support (Fig. 6.32). As found by W.H. and W.L. Bragg, diffracted X-rays beams result from reflection of X-rays by parallel planes of the same atoms or ions in a crystal (Fig. 6.33). The condition for constructive interference requires that the path difference between X-ray waves reflected by successive planes of atoms be a multiple of the wavelength,

$$2d \sin \theta = n\lambda \tag{6.34}$$

where d is the spacing between adjacent planes of atoms, θ the angle of incidence, n the diffraction order, and λ the X-ray wavelength.

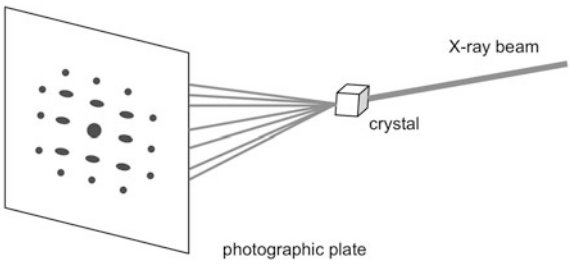


Fig. 6.32 X-ray diffraction experiment (schematic)

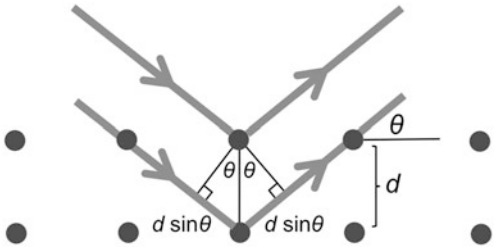


Fig. 6.33 Bragg diffraction

Each set of parallel planes of the same atoms or ions in the crystal acts like one three-dimensional diffraction grating. The diffracted X-ray intensity is mathematically related to the electronic density of the atoms by Fourier transforms. Solving a diffraction pattern consists in determining the crystalline structure that leads to the obtained diffraction pattern. This is a complex and time-consuming task that requires the expertise of an X-ray crystallographer, sophisticated software, and adequate computational resources.

In order to clarify the relationship between the diffraction pattern and the crystalline structure on purely qualitative and empirical grounds, a change of scale can be applied to an X-ray diffraction experiment, with the X-ray beam being replaced by a laser beam in the visible, and the 3D crystal being substituted by ordered patterns printed on 2D projection slides. Experiments like this are called **optical transforms** of the X-ray diffraction experiment (Lisensky et al. 1991).

Optical transforms correspond to a change of scale by a factor of order 10^4 on the X-rays' wavelength ($\lambda_{\text{X-ray}} \approx 10^{-10}$ m, $\lambda_{\text{laser}} \approx 10^{-6}$ m) and a factor of order 10^6 on the motif dimensions (an atom diameter $\approx 10^{-10}$ m, a motif on the slide 10^{-4} m). The diffraction pattern of an optical transform is obtained on a screen about 20 m from the slide, which is about 20 cm from the laser. For the slide perpendicular to the incident laser beam, the condition for constructive interference (Fraunhofer diffraction, Fig. 6.34) is given by

$$d \sin \theta = n\lambda \tag{6.35}$$

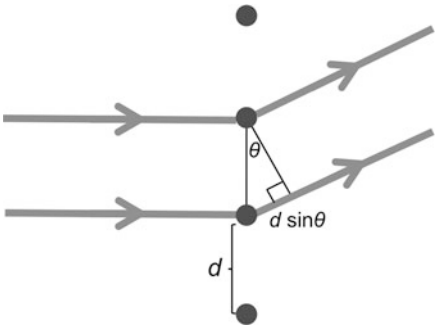


Fig. 6.34 Fraunhofer diffraction

At such a distance between slide and screen, the diffraction pattern is about 1° wide (1° = 0.0175 radians) and $\sin\theta \approx \theta$. Considering first-order diffraction ($n = 1$), (6.35) can be replaced by

$$d\,\theta \approx \lambda \tag{6.36}$$

in the optical transform. This expression allows one to correlate changes in the ordered pattern on the slide (changes in d) with changes produced in the optical transform (changes in θ). For the same λ, d and θ are inversely related.

The method of optical transforms has been used to show that B-DNA structural parameters such as the spacing between base pairs and the radius of the vertical section of the backbone structure can be obtained from the optical transform of the B-DNA X-ray diffraction pattern (Lucas et al. 1999).

6.5 Electrons in Solids

Electrons in atoms and molecules have discrete energies and are said to occupy defined **energy levels**, since such can be confirmed by spectroscopic methods. However, in a solid, the proximity between an atom and its neighbors leads to significant overlap between electronic wave functions and to electron delocalization, resulting in large numbers of energy levels with closely spaced energies, which in practical terms, form a continuum of energy levels, or **band**.

Consider a planar hydrocarbon with CC bonds of uniform bond distance and uniform bond order (1.5) and use the Hückel empirical method to calculate the energy levels of the system of π electrons as the number of carbon atoms increases (McWeeny 1979). In this model system, each carbon atom has one electron in one p orbital perpendicular to the backbone of carbon atoms. The overlap between adjacent p orbitals leads to π electron delocalization over the entire carbon backbone (Fig. 6.35).

In its original form, Hückel's method was conceived to deal with delocalized π electron systems and uses two parameters, α and β, which represent the ionization energy of a carbon atom and the interaction energy between p_π orbitals of adjacent

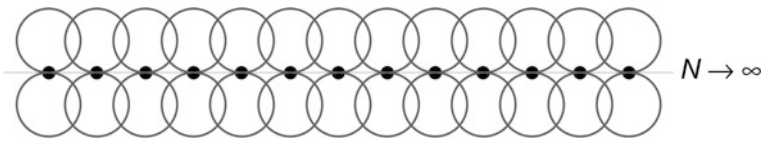


Fig. 6.35 The π electron system of a linear polyene hydrocarbon can inspire a model for a unidimensional crystal. Figure obtained with *Mathematica*

carbon atoms, respectively. The parameters α and β are negative quantities. For a linear polyene, the k molecular orbital energy calculated by Hückel's method is given by

$$\varepsilon_k = \alpha + 2\beta \cos \frac{\pi}{N+1} k \tag{6.37}$$

where N is the number of carbon atoms. For $k = 1$ (the lowest energy level) and $N \to \infty$,

$$\lim_{N\to\infty} \varepsilon_1 = \alpha + 2\beta \tag{6.38}$$

In turn, for $k = N$ (the highest energy level) and $N \to \infty$,

$$\lim_{N\to\infty} \varepsilon_N = \alpha - 2\beta \tag{6.39}$$

Making use of (6.37), we see that the energy difference between consecutive levels is given by

$$\begin{aligned} \Delta\varepsilon = \varepsilon_{k+1} - \varepsilon_k &= 2\beta\left(\cos\frac{\pi(k+1)}{N+1} - \cos\frac{\pi k}{N+1}\right) \\ &= -4\beta \sin\frac{\pi}{2(N+1)} \sin\frac{(2k+1)\pi}{2(N+1)} \end{aligned} \tag{6.40}$$

(**E9**).

With a linear polyene with $N \to \infty$ taken as a model for a unidimensional crystal, the energy levels are limited by the lowest and highest energies [(6.38) and (6.39)] that do not depend on k. In addition, from (6.40), it can be concluded that the energy difference between consecutive levels decreases as N increases. Therefore, for $N \to \infty$, the energy difference (6.40) tends to zero, and the energy levels are replaced by a continuum of levels called a **band** (Fig. 6.36). The *Mathematica* code **M11** draws this figure.

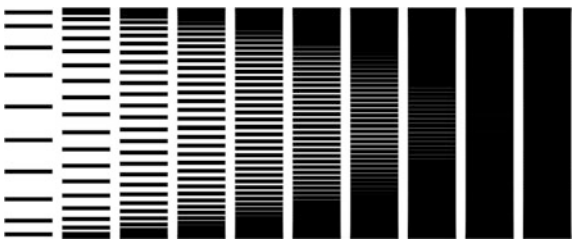


Fig. 6.36 Energy levels for the π electron delocalized system of a linear polyene, calculated by Hückel's method, as the number of carbon atoms increases from 10 to 100, in steps of 10. Figure obtained with *Mathematica*

6.6 Semiconductors

Experimental information on the electronic structure of solids can be obtained from values of electrical and thermal properties. In particular, the electrical resistivity and its thermal behavior can lead to a general classification of materials and details of their electronic structure. Electrical resistance is defined by $R \equiv V/I$, where V and I represent an electric potential difference and an electric current intensity, respectively. The variable R is an extensive variable. For a cylindrical material of length L and uniform sectional area S, one has $R = \rho\, L/S$, where ρ is the **electrical resistivity** (ρ is an intensive variable and its inverse is the electrical conductivity $\sigma = 1/\rho$) and the thermal coefficient of electrical resistivity is given by

$$\alpha = \frac{1}{\rho}\frac{d\rho}{dT} \tag{6.41}$$

with units K^{-1} or $°C^{-1}$. Typical values for the resistivity of metals are of order of magnitude 10^{-8} Ω m, and they increase with temperature. In particular, the resistivities of iron, beryllium, gold, and copper increase almost linearly with temperature, except at very low temperatures, where the increase is exponential (Fig. 6.37). Solids with resistivity of this order of magnitude that increase with temperature are called metallic **conductors**. With high values of electrical resistivity, we find materials called electrical **insulators**, such as glasses with electrical resistivities 10^{10}–10^{14} Ω m and polytetrafluoroethylene (PTFE, trademark Teflon) with resistivities 10^{22}–10^{24} Ω m. Materials with intermediate values of electrical resistivities (10^{-5}–10 Ω m) are called **semiconductors**. Semiconductors have the

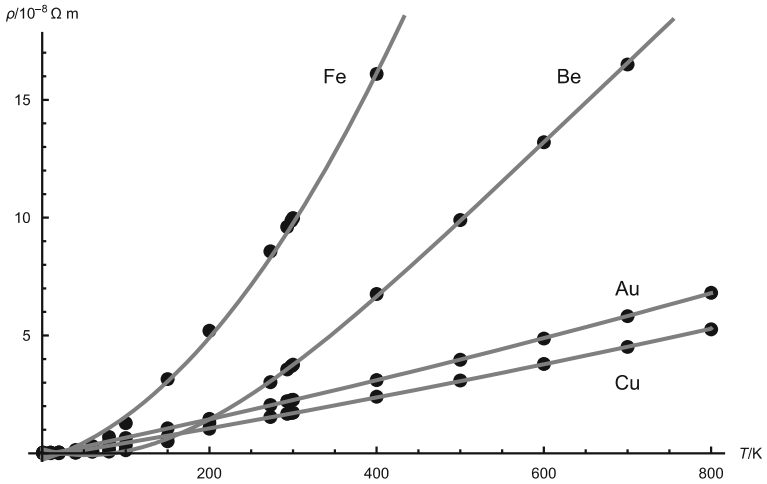


Fig. 6.37 Resistivity as a function of temperature for iron, beryllium, gold, and copper. Data values taken from Handbook of Chemistry and Physics, 2011 and figure obtained with *Mathematica*

electric current transported by electrons, as occurs with metals, but unlike metals, they exhibit an electrical resistivity that decreases exponentially with temperature when they are extremely pure, and the resistivity may even decrease by several orders of magnitude when they are doped with trace quantities of an electrical active material.

The valence band of insulators is fully occupied with valence electrons, and the band gap between the valence band and the lowest energy unoccupied band is far greater than the energy of the medium (energy associated with the thermal motion of atoms, sometimes referred to as thermal energy as its name suggests, it is zero at $T = 0$ K). Thus, electrons of the highest energy in the valence band, that is, at the band top, do not have unoccupied neighboring levels available to be excited. Materials with this band structure are **insulators**.

Unlike the valence band of electrical insulators, the valence band of **metallic conductors** is not fully occupied, and a temperature increase causes more collisions between excited electrons and atoms, thus leading to an increased resistivity. The mean number of electrons in each energy level is given by the **Fermi–Dirac distribution**, a variant of the Boltzmann distribution that takes into account the Pauli exclusion principle. At a defined temperature T, i.e., at thermal equilibrium, the mean number of electrons in energy level ε is given by

$$\phi(\varepsilon) = \frac{1}{1 + \exp\left[(\varepsilon - \mu) / (k_B T)\right]} \tag{6.42}$$

where k_B is the Boltzmann constant and μ is such that when $\varepsilon = \mu$, then $\phi = 1/2$, i.e., 50 % of the energy level with energy μ is occupied. The inverse function of (6.42) that gives the energy ε as a function of the population of electrons ϕ is

$$\varepsilon = \mu + k_B T \ln\left(\frac{1}{\phi} - 1\right) \tag{6.43}$$

This function is plotted in Fig. 6.38 for sodium ($\mu = 3.24$ eV), at 100, 1200, 4000, and 10,000 K (see *Mathematica* code **M12**, **E10**).

Semiconductors have a band structure qualitatively similar to that of electrical insulators, with their valence band being fully occupied. However, unlike insulators, the band gap for **semiconductors** is small and can be overcome by thermal excitation of electrons at temperatures greater than a minimum value defined by the band gap [Fig. 6.39a]. The excitation of electrons from the valence band leaves the electron holes behind. In the presence of an electric field, excited electrons and electron holes move in opposite directions. In fact, an electron moves in the opposite direction to the applied electric field, whereas an electron hole moves in the direction of the applied electric field, thus behaving like a positive charge. In a semiconductor, both the electrons of the conduction band and the holes in the valence band contribute to the electric current.

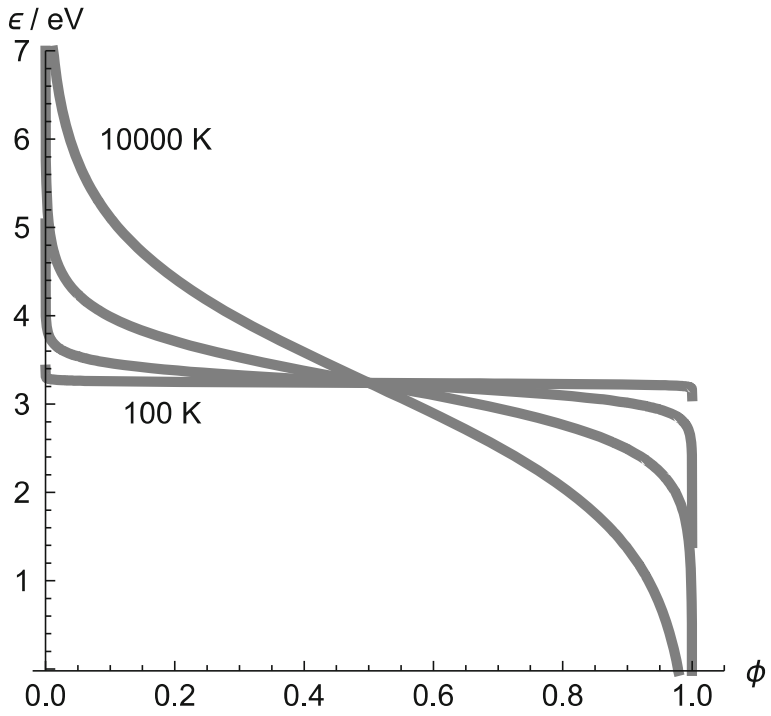


Fig. 6.38 The Fermi–Dirac distribution for sodium (μ = 3.24 eV) at T = 100, 1200, 4000, and 10,000 K shown as energy as a function of electrons population. Graph obtained with *Mathematica*

Consider now a silicon crystal and the substitution of a few of its atoms by arsenic in a mixing proportion of about one arsenic atom per one million atoms of silicon (**doping** of silicon with arsenic). Arsenic has one more valence electron than silicon. Thus, one arsenic atom is equivalent to a positive ion, isoelectronic with a silicon atom, plus one electron added to the silicon lattice. At such a small mixing proportion, it is unlikely that the crystalline network of silicon will be significantly modified or disrupted, an undesirable effect that would cause an abrupt increase in electrical resistivity. The addition of one electron per arsenic atom to the silicon lattice results in a small **donor band**, just below the unoccupied band [Fig. 6.39b], with an energy gap smaller than the medium thermal energy Therefore, thermal excitation of electrons to the unoccupied band becomes possible, and the electrical resistivity decreases with temperature (***n*-type semiconductor**).

If the silicon crystal is doped with boron, whose atoms have one valence electron fewer than silicon, then each boron atom is equivalent to a negative ion, isoelectronic with silicon, plus one electron hole added to the silicon lattice that acts like a positive charge. Formation of electron holes leads to charge separation (electrons and holes move in opposite directions). Since energy is required to separate charges, the acceptor band forms just above the fully occupied valence band [Fig. 6.39b].

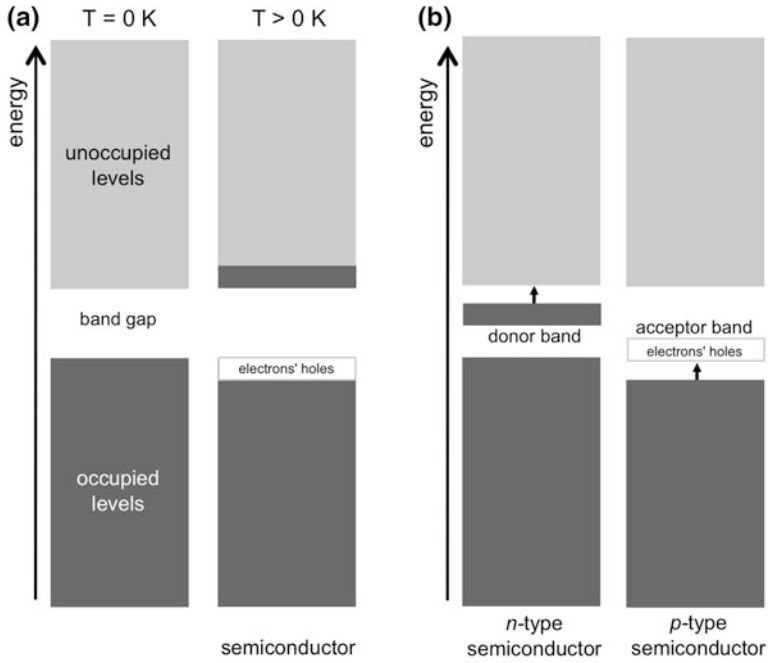


Fig. 6.39 Valence bands for semiconductors (schematic). **a** Valence band for pure semiconductor at T = 0 K and T > 0 K. **b** Valence bands for n-type and p-type semiconductors

Therefore, electrons from the valence band are excited to the acceptor band using the medium thermal energy and combine with holes in the acceptor band, causing a decrease of electrical resistivity with temperature (***p*-type semiconductor**).

Mathematica Codes

M1. Packing of Disks

The following *Mathematica* code builds disk lattices for square packing, whereby each disk is tangent to four neighboring disks, and hexagonal packing, whereby each disk is tangent to six neighboring disks. The *Mathematica* graphics primitive for a disk is

```
Disk[{x,y},r]
```

where `x` and `y` are the Cartesian coordinates of the center of a disk with radius `r`. For a square lattice of disks with radius 1/2, each of the Cartesian coordinates of the disk centers varies in one-unit increments, and so the array of coordinates are given by

```
Table[Disk[{i,j},1/2],{i,3},{j,3}]
```

where `i` and `j` correspond to `x` and `y`, respectively. Hence, a fixed `x` and variable `y` correspond to a row of disks, whereas a fixed `y` and variable `x` give a column of

disks. Note that the outermost `j` of the `Table` function changes more rapidly than the innermost `i`. This means that the disks are organized by rows, and the coordinates of the disk centers in the square packing are given by

```
Table[{i,j},{i,3},{j,3}]
```

`EdgeForm[Black]` is a graphics directive that specifies that the edge of the disks should be drawn in `Black`. In turn, the graphics directive `LightGray` indicates that the disks should be filled in this color.

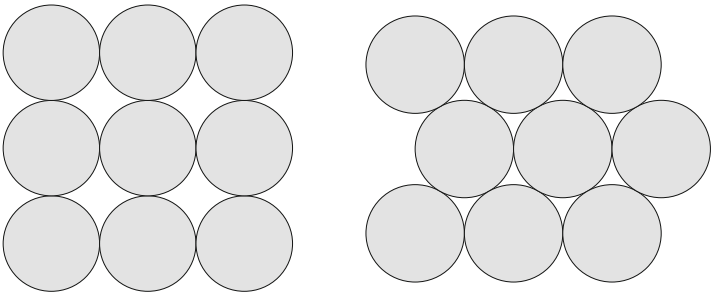
For a hexagonal close packing columns with odd `j` have `x=i`, and columns with even `j` have `x=i+1/2`. In the *Mathematica* code, we use `Boole[EvenQ[j]]`, which gives the result 0 if `EvenQ[j]` is `False` (`j` is odd) and produces 1 if `EvenQ[j]` is `True` (`j` is even). In addition, by inspection of the triangle of disks in Fig. 6.1, it can be found that `y=(√3/2)j`. Hence, the hexagonal close packing of disks can be obtained with

```
Table[Disk[{i+Boole[EvenQ[j]]/2,Sqrt[3]/2 j},1/2],{i,3},{j,3}]
```

and the coordinates of disk centers are given by

```
Table[{i+Boole[EvenQ[j]]/2,Sqrt[3]/2 j},{i,3},{j,3}]
```

```
aa=Graphics[{EdgeForm[Black],LightGray,Table[Disk[{i,j},1/2],{i,3},{j,3}]},
   ImageSize→300];
bb=Graphics[{EdgeForm[Black],LightGray,Table[Disk[{i+Boole[EvenQ[j]]/2,
   Sqrt[3]/2 j},1/2],{i,3},{j,3}]},ImageSize→300];
Row[{aa,"      ",bb}]
rr=Table[{i,j},{i,3},{j,3}]//MatrixForm
ss=Table[{i+Boole[EvenQ[j]]/2,Sqrt[3]/2 j},{i,3},{j,3}]//MatrixForm
Row[{rr,"      ",ss}]
```



$$\begin{pmatrix} \begin{pmatrix}1\\1\end{pmatrix} & \begin{pmatrix}1\\2\end{pmatrix} & \begin{pmatrix}1\\3\end{pmatrix} \\ \begin{pmatrix}2\\1\end{pmatrix} & \begin{pmatrix}2\\2\end{pmatrix} & \begin{pmatrix}2\\3\end{pmatrix} \\ \begin{pmatrix}3\\1\end{pmatrix} & \begin{pmatrix}3\\2\end{pmatrix} & \begin{pmatrix}3\\3\end{pmatrix} \end{pmatrix}$$

$$\begin{pmatrix} \begin{pmatrix}1\\\frac{\sqrt{3}}{2}\end{pmatrix} & \begin{pmatrix}\frac{3}{2}\\\sqrt{3}\end{pmatrix} & \begin{pmatrix}1\\\frac{3\sqrt{3}}{2}\end{pmatrix} \\ \begin{pmatrix}2\\\frac{\sqrt{3}}{2}\end{pmatrix} & \begin{pmatrix}\frac{5}{2}\\\sqrt{3}\end{pmatrix} & \begin{pmatrix}2\\\frac{3\sqrt{3}}{2}\end{pmatrix} \\ \begin{pmatrix}3\\\frac{\sqrt{3}}{2}\end{pmatrix} & \begin{pmatrix}\frac{7}{2}\\\sqrt{3}\end{pmatrix} & \begin{pmatrix}3\\\frac{3\sqrt{3}}{2}\end{pmatrix} \end{pmatrix}$$

M2. Hexagon of Disks

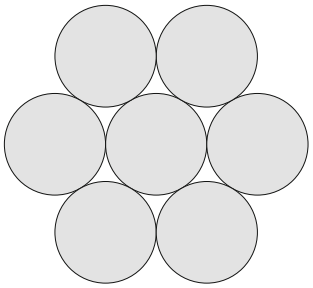
The close packing of disks in a plane shows that each disk is tangent to six disks, whose centers are the vertices of a regular hexagon. The following *Mathematica* code presents a hexagonal arrangement of disks around a central disk, by taking advantage of its sixfold rotational symmetry axis. The *Mathematica* code that leads to this arrangement is

```
Graphics[{EdgeForm[Black],Lightgray,Disk[{0,0},1/2],
  Table[{EdgeForm[Black],Lightgray,Disk[{Cos[2πi/6],
  Sin[2πi/6],1/2]},{i,0,5}]
```

The coordinates of the disks centers are given by

```
{{0,0},Table[{Cos[2πi/6],Sin[2πi/6]},{i,0,5}]
```

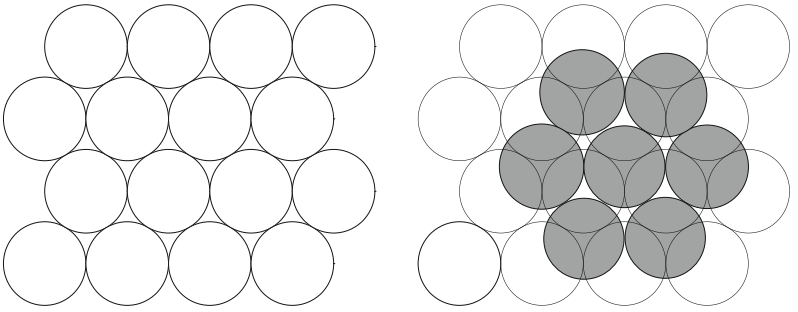
```
Graphics[{EdgeForm[Black],LightGray,Disk[{0,0},1/2],
   Table[{EdgeForm[Black],LightGray,Disk[{Cos[2πi/6],Sin[2πi/6]},1/2]},
   {i,0,5}]},ImageSize→300]
{{0,0},Table[{Cos[2πi/6],Sin[2πi/6]},{i,0,5}]}
```



$$\left\{\{0,0\},\left\{\{1,0\},\left\{\frac{1}{2},\frac{\sqrt{3}}{2}\right\},\left\{-\frac{1}{2},\frac{\sqrt{3}}{2}\right\},\{-1,0\},\left\{-\frac{1}{2},-\frac{\sqrt{3}}{2}\right\},\left\{\frac{1}{2},-\frac{\sqrt{3}}{2}\right\}\right\}\right\}$$

M3. Disk Layers

```
aa=Graphics[{EdgeForm[Black],White,Table[Disk[{i+Boole[EvenQ[j]]/2,
   Sqrt[3]/2 j},1/2],{i,4},{j,4}]},ImageSize→300];
bb=Graphics[{EdgeForm[Black],Opacity[0.5,Gray],Disk[{3,7Sqrt[3]/6},1/2],
   Table[{EdgeForm[Black],Opacity[0.5,Gray],Disk[{Cos[2πi/6]+3,
   Sin[2π i/6]+7Sqrt[3]/6},1/2]},{i,0,5}]},ImageSize→300];
cc=Show[aa,bb];
Row[{aa,"   ",cc}]
```



In the picture on the right, this *Mathematica* code considers two close-packed layers of identical disks, where the second layer is displaced so that its disk centers coincide with the interstices of the first layer. The translation vector that shifts the second layer of disks with respect to the first layer is shown in Fig. 6.4, where the coordinates of its origin and its tip at the center of the interstice are $(1, \sqrt{3}/2)$ and $(3/2, 2/\sqrt{3})$, respectively. Therefore, the translation vector is given by

$$\mathbf{t} = \frac{1}{2}\hat{\mathbf{x}} + \frac{\sqrt{3}}{6}\hat{\mathbf{y}}$$

[see (6.1)]. `Opacity` varies from 0 to 1, where 0 corresponds to perfect transparency, and `Show` presents several graphics combined.

Suggestion: Modify the above code so that it prints the coordinates of the disk centers shown in gray.

M4. The Third Dimension

The *Mathematica* `Graphics3D` primitive for a sphere is

```
Sphere[{x,y,z},r]
```

where `x`, `y`, `and z` are the Cartesian coordinates of the center of a sphere with radius `r`. In order to determine the `z` coordinate of a second horizontal layer of close-packed spheres, we begin by considering four equivalent and mutually tangent spheres of radius 1/2 whose centers form a regular tetrahedron inside a cube (see Fig. 6.6). The distance between center 4 and the plane 1-2-3 is given by $\sqrt{2}/\sqrt{3}$, as can be shown by simple trigonometry. Sphere 4 is nested on the interstice of spheres 1, 2, and 3. Therefore, $\sqrt{2}/\sqrt{3}$ is the vertical distance between two horizontal close-packed layers of spheres of radius 1/2. The vector that goes from center 1 to center 4 is equal to the sum of vectors (4-c) and (c-1) and is given by

$$\mathbf{t} = \frac{1}{2}\hat{\mathbf{x}} + \frac{\sqrt{3}}{6}\hat{\mathbf{y}} + \sqrt{\frac{2}{3}}\hat{\mathbf{z}}$$

The figure below on the right shows a regular tetrahedron formed by four mutually tangent spheres. The *Mathematica* line of code that leads to the arrangement of three spheres on the left takes advantage of the threefold rotational symmetry axis and is given by

```
Graphics3D[Table[{Gray,Sphere[{f*Sin[2πi/3],
  f*Cos[2πi/3],0.0,1/2]},{i,0,2}]
```

where

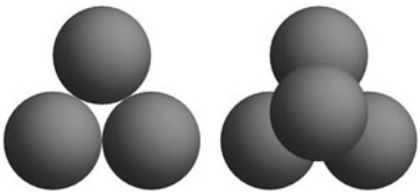
```
f=1/Sqrt[3]
```

and

`f*Sin[2πi/3]=x` and `f*Cos[2πi/3]=y`

(see Fig. 6). `Viewpoint→{0,0,Infinity}` specifies a plan view from above, supposedly from infinite distance.

```
aa=Graphics3D[Table[{Gray,Sphere[{Sin[2πi/3]/Sqrt[3],Cos[2πi/3]/Sqrt[3],0.0},
   1/2]},{i,0,2}],Lighting→"Neutral",Boxed→False,ViewPoint→{0,0,Infinity},
   ImageSize→300];
bb=Graphics3D[{Table[{Gray,Sphere[{Sin[2πi/3]/Sqrt[3],Cos[2πi/3]/Sqrt[3],0.0},
   1/2]},{i,0,2}],Gray,Sphere[{0.,0.,Sqrt[2/3]},1/2],Lighting→"Neutral",Boxed→False,
   ViewPoint→{0,0,Infinity},ImageSize→300];
Row[{aa," ",bb}]
{Table[{Sin[2πi/3]/Sqrt[3],Cos[2πi/3]/Sqrt[3],0.0},{i,0,2}],{0.0,0.0,Sqrt[2/3]}}
```

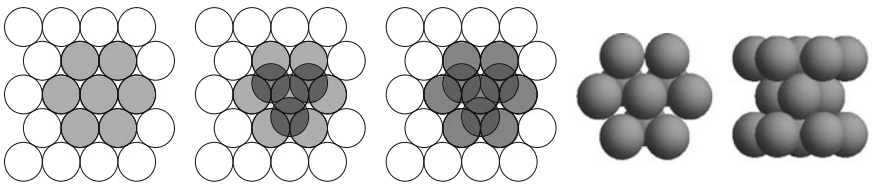


$$\left\{\left\{\left\{0,\ \frac{1}{\sqrt{3}},\ 0.\right\},\ \left\{\frac{1}{2},\ -\frac{1}{2\sqrt{3}},\ 0.\right\},\ \left\{-\frac{1}{2},\ -\frac{1}{2\sqrt{3}},\ 0.\right\}\right\},\ \left\{0.,\ 0.,\ \sqrt{\frac{2}{3}}\right\}\right\}$$

M5. HCP Structure

The following *Mathematica* code draws a figure for the hexagonal close packing of disks and rigid spheres with equal radius:

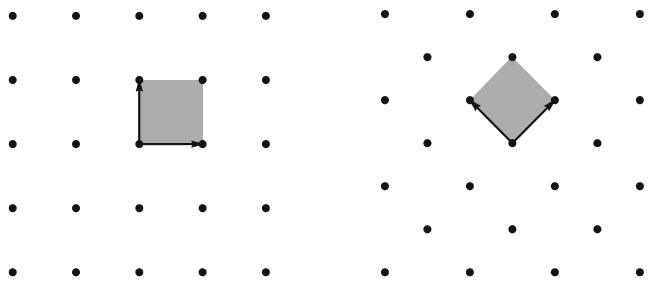
```
d=1/Sqrt[3];
(* 2D *)
aa=Graphics[{EdgeForm[Black],White,Table[Disk[{i+Boole[EvenQ[j]]/2,
   Sqrt[3]/2 j},1/2],{i,4},{j,5}]},ImageSize→250];
bb=Graphics[{EdgeForm[Black],Opacity[0.5,LightGray],Disk[{3,3Sqrt[3]/2},1/2],
   EdgeForm[Black],Opacity[0.5,LightGray],Table[Disk[{Cos[2πi/6]+3,
   Sin[2πi/6]+3Sqrt[3]/2},1/2],{i,0,5}]}];
cc=Graphics[{EdgeForm[Black],Opacity[0.5,Darker[Gray]],Table[Disk[{-dSin[2πi/3]+3,
   -dCos[2πi/3]+3Sqrt[3]/2},1/2],{i,0,2}]}];
a1=Show[aa,bb];
a2=Show[aa,bb,cc];
a3=Show[aa,bb,cc,bb];
(* 3D *)
dd=Graphics3D[{Gray,Sphere[{3,3 Sqrt[3]/2,-Sqrt[2/3]},1/2],Gray,
   Table[Sphere[{Cos[2πi/6]+3,Sin[2πi/6]+3Sqrt[3]/2,-Sqrt[2/3]},1/2],{i,0,5}]},
   Lighting→"Neutral",Boxed→False,ImageSize→250];
ee=Graphics3D[{Gray,Table[Sphere[{-dSin[2πi/3]+3,-dCos[2πi/3]+3Sqrt[3]/2,
   0.0},1/2],{i,0,2}]},Lighting→"Neutral"];
ff=Graphics3D[{Gray,Sphere[{3,3Sqrt[3]/2,Sqrt[2/3]},1/2],Gray,
   Table[Sphere[{Cos[2πi/6]+3,Sin[2πi/6]+3Sqrt[3]/2,Sqrt[2/3]},1/2],{i,0,5}]},
   Lighting→"Neutral"];
b1=Show[dd,ee,ff,ViewPoint→{0,0,Infinity}];
b2=Show[dd,ee,ff,ViewPoint→{0,-Infinity,0}];
Row[{a1," ",a2," ",a3," ",b1," ",b2}]
```



M6. 2D Square Point Lattice

The following *Mathematica* code uses (6.12) and (6.13) to build a square point lattice to which is applied a π/4 anticlockwise rotation (`RotationMatrix` rotates 2D vectors anticlockwise). Each point on the lattice corresponds to a defined pair of integers `{n1,n2}`. Both arrays of points represent the same lattice, since the translation vectors `w1` and `w2` result from `v1` and `v2` by a π/4 counterclockwise rotation [see (6.15)]. The *Mathematica* command

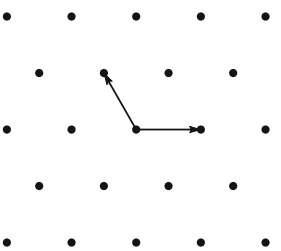
```
v1={1,0}; v2={0,1};
coords1=Table[n1 v1+n2 v2,{n1,-2,2},{n2,-2,2}];
latticeA=Flatten[coords1,1];
aa=Graphics{{Opacity[0.4],Gray,Polygon[{{0,0},v1,v1+v2,v2}]},
   {Black,Thick,Arrow[{{0,0},v1}]},{Black,Thick,Arrow[{{0,0},v2}]},
   {Black,PointSize[0.03],Point[latticeA]}},PlotRange→{{-2.1,2.1},{-2.1,2.1}},
   ImageSize→300,Axes→False];
r=RotationMatrix[45 Degree];
w1=r.v1; w2=r.v2;
coords2=Table[n1 w1+n2 w2,{n1,-3,3},{n2,-3,3}];
latticeB=Flatten[coords2,1];
bb=Graphics{{Opacity[0.4],Gray,Polygon[{{0,0},w1,w1+w2,w2}]},
   {Black,Thick,Arrow[{{0,0},w1}]},{Black,Thick,Arrow[{{0,0},w2}]},
   {Black,PointSize[0.029],Point[latticeB]}},PlotRange→{{-2.3,2.3},{-2.3,2.3}},
   ImageSize→310,Axes→False];
figure=Row[{aa,"          ",bb}]
```



M7. 2D Hexagonal Point Lattice

The following *Mathematica* code builds a hexagonal point lattice.

```
v1={1,0}; v2={-1/2,Sqrt[3]/2};
Graphics[{{Black,Thick,Arrow[{{0,0},v1}]},{Black,Thick,Arrow[{{0,0},v2}]},{Black,
   {PointSize[0.03],Point[Flatten[Table[n1 v1+n2 v2,{n1,-3,3},
   {n2,-3,3}],1]]}},PlotRange→{{-2.1,2.1},{-2.1,2.1}},ImageSize→300,
   Axes→False]
```

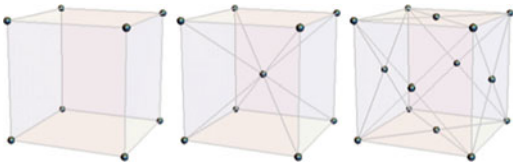


M8. 3D Bravais Lattices

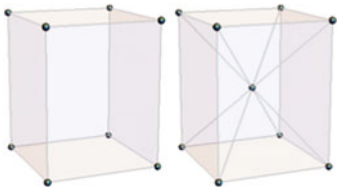
The following code uses the built-in *Mathematica* command `LatticeData` to present images for the 14 Bravais lattices corresponding to 7 crystallographic systems. Note that this command may require Internet connectivity:

```
c1=LatticeData["SimpleCubic",Image];
c2=LatticeData["BodyCenteredCubic",Image];
c3=LatticeData["FaceCenteredCubic",Image];
t1=LatticeData["SimpleTetragonal",Image];
t2=LatticeData["CenteredTetragonal",Image];
o1=LatticeData["SimpleOrthorhombic",Image];
o2=LatticeData["BaseCenteredOrthorhombic",Image];
o3=LatticeData["BodyCenteredOrthorhombic",Image];
o4=LatticeData["FaceCenteredOrthorhombic",Image];
m1=LatticeData["SimpleMonoclinic",Image];
m2=LatticeData["BaseCenteredMonoclinic",Image];
tc=LatticeData["SimpleTriclinic",Image];
tr=LatticeData["SimpleTrigonal",Image];
hx=LatticeData["SimpleHexagonal",Image];
Print["Cubic: Simple, BodyCentred, FaceCentred"]
aa=Row[{c1,c2,c3}]
Print["Tetragonal: Simple, Centred"]
bb=Row[{t1,t2}]
Print["Orthorhombic: Simple, BaseCentred, BodyCentred, FaceCentred"]
cc=Row[{o1,o2,o3,o4}]
Print["Monoclinic: Simple, BaseCentred"]
dd=Row[{m1,m2}]
Print["SimpleTriclinic"]
ee=Row[{tc},ImageSize→120]
Print["SimpleTrigonal"]
ff=Row[{tr},ImageSize→120]
Print["SimpleHexagonal"]
gg=Row[{hx},ImageSize→120]
```

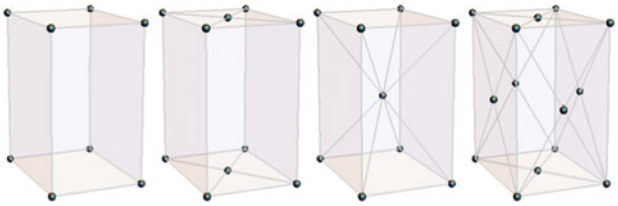
```
Cubic: Simple, BodyCentred, FaceCentred
```



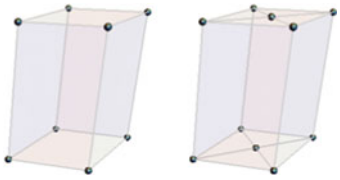
```
Tetragonal: Simple, Centred
```



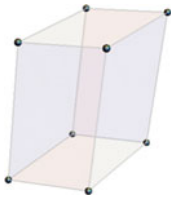
```
Orthorhombic: Simple, BaseCentred, BodyCentred, FaceCentred
```



```
Monoclinic: Simple, BaseCentred
```

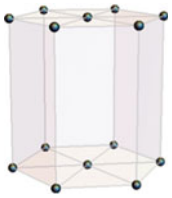


SimpleTriclinic



SimpleTrigonal

SimpleHexagonal



M9. CsCl Structure

The following *Mathematica* code builds the cesium chloride structure with the Cs^+ and Cl^- ions represented by spheres. If one Cl^- is at the origin, the translation vectors for the chloride ion lattice are given by (6.23) (Fig. 6.23) (simple cubic lattice), and the vector that translates this lattice to produce the cesium ion lattice is given by (6.29). The code forms two lattices, one for Cs^+, the other for Cl^-. One of the most important *Mathematica* commands in this code is `Cases`, which gives a list of elements that match a pattern. The output of `Cases` is a list of `True` and `False` values. Another important command is `Pick`, which selects elements of a list of `True` and `False` values. To illustrate the use of these two *Mathematica* commands, we consider the following line,

```
Cases[{{-4.12},{-6.18},{2.06},{6.18}},{x_}→x<5&&x>-5]
```

whose output is

```
{True,False,True,False}
```

and

```
Pick[{aaa,bbb,ccc,ddd},{True,False,True,False}]
```

whose output is

```
{aaa,ccc}
```

Note that the above two lines of code can be condensed into the following single line:

```
Pick[{aaa,bbb,ccc,ddd},Cases[{{-4.12},{-6.18},{2.06},
{6.18}},{x_}→x<5&&x>-5]]
```

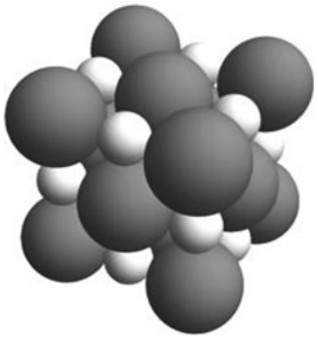
The following *Mathematica* code uses `Pick` and `Cases` combined in the same lines of code to define `latticeCs` and `latticeCl`:

```
a=4.12; rCs=1.69; rCl=1.81;
(* values in angstrom taken from Ashcroft and Mermin (see Further Reading) *)
v1=a{1,0,0}; v2=a{0,1,0}; v3=a{0,0,1}; t=a/2{1,1,1};
listCl=Flatten[Table[n1 v1+n2 v2+n3 v3,{n1,-2,2},{n2,-2,2},{n3,-2,2}],2];
latticeCl=Pick[listCl,Cases[listCl,{x_,y_,z_}→(x<5&&x>-5)&&(y<5&&y>-5)&&
    (z<5&&z>-5)]];
listCs=Flatten[Table[n1 v1+n2 v2+n3 v3+t,{n1,-2,2},{n2,-2,2},{n3,-2,2}],2];
latticeCs=Pick[listCs,Cases[listCs,{x_,y_,z_}→(x<5&&x>-5)&&(y<5&&y>-5)&&
    (z<5&&z>-5)]];
scCl=Graphics3D[{Opacity[1],Gray,Sphere[latticeCl,rCl]},Lighting→"Neutral",Axes→False];
scCs=Graphics3D[{Opacity[1],LightGray,Sphere[latticeCs,rCs]},Lighting→"Neutral"];
Show[{scCl,scCs},Boxed→False,ImageSize→250,ViewPoint→{0.8,-2.4,2}]
```

M10. NaCl Structure

This *Mathematica* code builds the sodium chloride structure with the Na^+ and Cl^- ions represented by spheres. This structure results from the interpenetration of two face-centered cubic structures, one for the chloride ions, the other for the sodium ions. The translation vectors for the face-centered cubic structure are given by (6.25). If one Cl^- is at the origin, the vector that translates the lattice of Cl^- ions to generate the sodium ion lattice is given by (6.30):

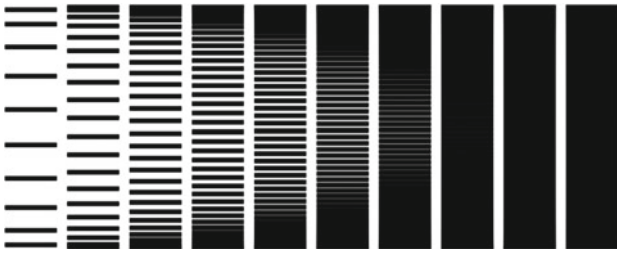
```
a=5.64; rNa=0.95; rCl=1.81; (* Ashcroft and Mermin *)
v1=(a/2){0,1,1}; v2=(a/2){1,0,1}; v3=(a/2){1,1,0}; t=(a/2){0,0,1};
listNa=Flatten[Table[n1 v1+n2 v2+n3 v3,{n1,-2,2},{n2,-2,2},{n3,-2,2}],2];
latticeNa=Pick[listNa,Cases[listNa,{x_,y_,z_}→(x<5&&x>-5)&&(y<5&&y>-5)&&(z<5&&z>-5)]];
listCl=Flatten[Table[n1 v1+n2 v2+n3 v3+t,{n1,-2,2},{n2,-2,2},{n3,-2,2}],2];
latticeCl=Pick[listCl,Cases[listCl,{x_,y_,z_}→(x<5&&x>-5)&&(y<5&&y>-5)&&(z<5&&z>-5)]];
fccCl=Graphics3D[{Opacity[1],Gray,Sphere[latticeCl,rCl]},Lighting→"Neutral",Axes→False];
fccNa=Graphics3D[{Opacity[1],LightGray,Sphere[latticeNa,rNa]},Lighting→"Neutral"];
Show[{fccCl,fccNa},Boxed→False,ImageSize→250,ImageSize→{1.3,-2.4,2}]
```



M11. from Energy Levels to Bands

This *Mathematica* code draws energy levels for the π electron delocalized system of a linear polyene, calculated by Hückel's method [see (6.37)], as the number of carbon atoms increases from 10 to 100, in steps of 10:

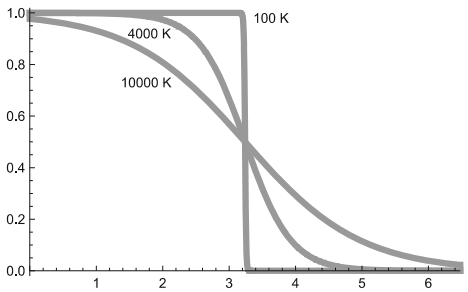
```
Row[Table[Plot[Table[N[2Cos[π k/(n+1)]],{k,1,n}],{x,0.12,0.88},
    Axes→{False,False},PlotStyle→{Thick,Black},PlotRange→{{0,1},{-2,2}},
    AspectRatio→4],{n,{10,20,30,40,50,60,70,80,90,100}}]]
```



M12. Fermi–Dirac Distribution

The following *Mathematica* code plots the Fermi–Dirac distribution (6.42) for sodium (μ = 3.24 eV) at T = 100, 4000, and 10,000 K, showing the electron population as a function of energy:

```
kB=1.3806488*10^-23/(1.6021766*10^-19); (* kB / eV K^-1 *)
μ = 3.24; (* see Ashcroft and Mermin, Further Reading *)
f[ε_,T_]:=1/(1+Exp[(ε-μ)/kB/T])
fermidirac=Plot[{f[ε,100],f[ε,4000],f[ε,10000]},{ε,0,2μ},
    PlotRange→{{-0.001,6.48},{-0.02,1.02}},Axes→True,
    PlotStyle→{{Gray,Thickness[0.015]},{Gray,
    Thickness[0.015]},{Gray,Thickness[0.015]}},
    ImageSize→300];
labels=Graphics[{Text[Style["100 K",FontFamily→"Arial",FontSize→9],{3.63,0.98}],
    Text[Style["4000 K",FontFamily→"Arial",FontSize→9],{1.8,0.92}],
    Text[Style["10000 K",FontFamily→"Arial",FontSize→9],{1.75,0.73}]}];
Show[fermidirac,labels]
```



Glossary

Bravais lattice An infinite array of discrete points in three-dimensional space such that for every point P of the array defined by the positional vector **p**, it is possible to find a set of integers (may be positive, negative, or zero) (n_1, n_2, n_3) that obey the following vector equation $\mathbf{p} = n_1\mathbf{v}_1 + n_2\mathbf{v}_2 + n_3\mathbf{v}_3$, where $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are linearly independent translation vectors (noncollinear vectors) that span the lattice

Coordination number For a system of rigid spheres of the same radius is the number of closest neighbors of each sphere

Electronic band The range of energies an electron may have in a solid. The electronic band structure of a solid is formed by electronic bands and band gaps that are the ranges of energies the electron may not have

Fermi–Dirac distribution	Gives the probability of occupancy of energy levels by electrons. It is a variant of the Boltzmann distribution that takes into account the Pauli exclusion principle
Optical transform	A two-dimensional array of dots or other motifs on a slide that with a beam of visible laser light produces a diffraction pattern. With respect to an X-ray diffraction experiment, an optical transform corresponds to a reduction from 3D, on the real crystal, to 2D, on the slide, a change of scale on the X-ray's wavelength and on the motif's dimensions. The diffraction pattern of an optical transform is obtained on a screen about 20 m from the slide, which is about 20 cm from the laser
Packing density	Of rigid spheres is the ratio of the volume of all spheres contained in a repeating unit and the volume of the repeating unit. The packing density measures the efficiency of occupation of the three-dimensional space by rigid spheres
Semiconductor	Material with electrical resistivity intermediate between that of a metallic conductor and an insulator in which electric current is transported by electrons, as happens with metals, but unlike metals, exhibits an electrical resistivity that decreases exponentially with temperature when it is extremely pure and decreases by several orders of magnitude when it is doped with trace quantities of an electrically active material
Unit cell	The smallest parallelepiped that contains all structural and symmetric information to build a crystal lattice by translation, without overlapping itself or leaving voids

Exercises

E1. Change the numerical values of Fig. 6.6 so that the new values refer to identical spheres of radius r.

E2. Use *Mathematica* to draw three tangent disks of radius 1/2 whose centers form an equilateral triangle and print coordinates of the disks' centers. Discuss the results.

E3. Calculate the packing density for the hexagonal close-packed structure.

E4. Consider a two-dimensional structure formed by the merging of two Bravais point lattices of black and gray points, where each black point is the center of an equilateral triangle of gray points with the vertex pointing downward and each gray point is the center of an equilateral triangle of black points with the vertex pointing upward. Find the translation vector for converting one Bravais lattice into the other, and use *Mathematica* to represent the whole two-dimensional structure (black and gray points).

E5. Use *Mathematica* to represent a cube of spheres in the sc structure. Assume $r = 1/2$ and take the lattice constant a so that the spheres become tangent along the cube edge.

E6. Use *Mathematica* to represent a cube of spheres in the bcc structure. Assume $r = 1/2$ and take the lattice constant a so that the spheres become tangent along the cube main diagonal.

E7. Use *Mathematica* to represent a cube of spheres in the fcc structure. Assume $r = 1/2$ and take the lattice constant a so that the obtained structure is cubic close-packed.

E8. Use *Mathematica* to represent a hexagonal prism of spheres in the hcp structure. Assume $r = 1/2$ and take the lattice constants a and c so that the obtained structure is close-packed.

E9. Use *Mathematica* to arrive at the last equality of (6.40).

E10. Consider the Fermi–Dirac distribution function for metallic sodium at 7000 K. Calculate the energy in eV of the level whose occupation fraction is equal to 0.25.

References

Kapustinskii AF (1956) Lattice energy of ionic crystals. Q Rev Chem Soc 10:283–294

Lisensky GC, Kelly TF, Neu DR, Ellis AB (1991) Simulating diffraction experiments in introductory courses. J Chem Educ 68:91–96

Lucas AA, Lambin Ph, Mairesse R, Mathot M (1999) Revealing the backbone structure of B-DNA from laser optical simulations of its X-ray diffraction diagram. J Chem Educ 76:378–383

McWeeny R (ed) (1979) Coulson's Valence, 3rd edn. Oxford University Press, Oxford

Further Reading

Ashcroft NW, Mermin ND (1976) Solid state physics. Brooks/Cole Publishers Co.

Atkins PW (1998) Physical chemistry, 6th edn. W. H. Freeman and Company

Smart, L. and Gagan, M., editors, The Third Dimension, The Open University, Series The Molecular World, 2002

Water

7

Abstract

Water determines life on earth to an extent greater than any other substance. We know it in three physical states: as a liquid, in the oceans, seas, lakes, and rivers; as a solid, in lakes covered with ice, in glaciers, and in the North and South Poles; as a vapor, in the Earth's atmosphere. Despite having a simple molecular formula, water has unique structural and thermodynamic properties: it is denser in the liquid than in the solid states (ice floats in liquid water), and it melts when ice is subject to high pressure. In addition, the specific heat capacity of liquid water is about four times that of dry air, and its surface tension is about three times that of ethanol. The computer simulation of water by molecular dynamics requires large computational resources and is still is a topic of current research. In this chapter, we consider several properties of the water molecule including its geometry, enthalpy of formation, atomic charges, dipole moment, molecular orbitals, molecular vibrations, and hydrogen bond. The remaining sections of this chapter take the standpoint of thermodynamics and consider normal ice, liquid water, the phase diagram of water, water as a solvent, simple nonpolar solutes, ionic solutions, amphipathic molecules, acids and bases, and standard electrode potentials. The Mathematica codes presented at the end of the chapter (Representation of the Water Molecule, Electrostatic Potential Contours for a Dipole, Interactive Manipulation of Charge, Vector Field Streamlines, Luzar's Model, Micelle-Monomers Equilibrium, Critical Micelle Concentration, Weak Acid HA) are accompanied with detailed explanations and provide insight into the corresponding subjects. Following the Mathematica codes, we present a glossary of important scientific and technical terms on water and aqueous solutions and a list of exercises, whose complete answers can be found after the Appendix.

J.J.C. Teixeira-Dias, *Molecular Physical Chemistry*,
DOI 10.1007/978-3-319-41093-7_7

7.1 Molecular Geometry

The experimental values of the OH bond length and the HOH bond angle of a water molecule are $R_{OH} = 0.958$ Å and $\alpha_{HOH} = 104.4776°$ (CCCBDB), showing that H_2O is a **symmetric nonlinear triatomic molecule** (Fig. 7.1; see *Mathematica* code **M1**). The geometry optimization of a water molecule carried out by the system of programs *Gaussian* 09 at the B3LYP level of calculation with the basis set cc-pVTZ gives $R_{OH} = 0.9611$ Å and $\alpha_{HOH} = 104.5296°$ (Fig. 7.1).

Figure 7.2 shows the energy profiles for the scans of the bond length R_{OH} when α_{HOH} is 104.53°, and of the bond angle α_{HOH} when R_{OH} is 0.961 Å. Note that the HOH bond angle differs from the tetrahedral angle (109.47°) by approximately 5°, a value much smaller than the HOH bond angle difference with respect to the right angle, which is about 19°.

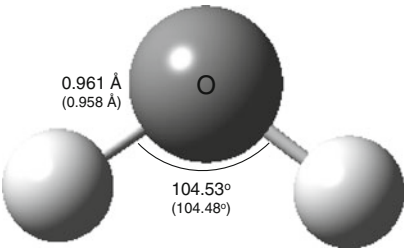


Fig. 7.1 Ball and bond representation of the H_2O molecule with calculated and experimental (CCCBDB) values of the geometric variables R_{OH} and α_{HOH} (experimental values are shown in *parentheses*). Molecular representation obtained with *GaussView*

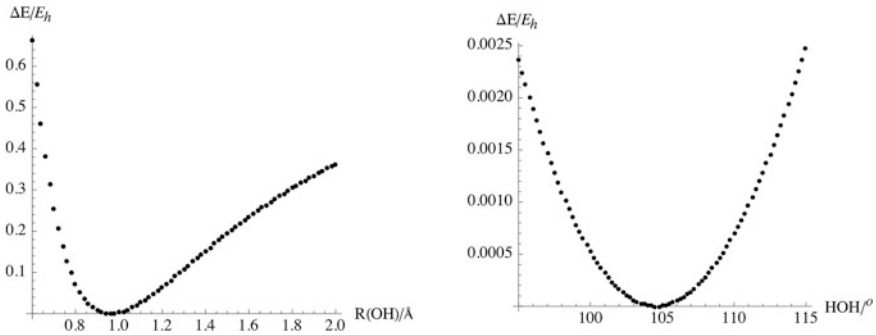


Fig. 7.2 At *left*, scan of the bond length R_{OH} when α_{HOH} is equal to 104.53°. At *right*, scan of the bond angle α_{HOH} with R_{OH} equal to 0.961 Å. Note that the energy scales of the above plots are quite different. Graphs obtained with *Mathematica*

7.2 Enthalpy of Formation

For a gas phase water molecule, the experimental value of the enthalpy of formation at 298.15 K is −241.81 kJ mol^{-1} (CCCBDB). Computing a molecular energy with chemical accuracy means attaining an accuracy of 1 kcal mol^{-1} (≈4.2 kJ mol^{-1}). Since the largest errors in ab initio calculations result from basis set truncation, achieving chemical accuracy requires a series of calculations that extrapolate to the complete basis set limit. This type of calculation was carried out by the system of programs *Gaussian* 09 using the composite method W1BD (Martin and de Oliveira 1999; Parthiban and Martin 2001), where W stands for Weizmann Institute and BD for Brueckner Doubles (Brueckner orbitals). Within W1BD, the geometry optimization is carried out by a B3LYP/cc-pVTZ calculation.

The Schrödinger equation is obtained within the Born–Oppenheimer approximation, and the electron wave function and its corresponding energy U_{elec} depend parametrically on the nuclear coordinates. The energy U_{elec} is evaluated at 0 K. In order to convert this energy from 0 to 298.15 K, the **thermal energy** given by the sum of the translation, rotation, and vibration contributions ($U_{\text{trans}} + U_{\text{rot}} + U_{\text{vib}}$) at 298.15 K needs to be added to U_{elec}. *Gaussian* evaluates the thermal energy and the enthalpy at 298.15 K (the enthalpy H is defined as $U + pV$, with the product pV being equal to RT for a mole of an ideal gas). In order to obtain the enthalpy of formation of $H_2O(g)$, the same composite calculation method has to be applied to H_2O, H_2, and O_2, since the chemical reaction for the formation of one mole of water in the gas phase is given by $H_2(g) + (1/2)\ O_2(g) \rightarrow H_2O(g)$. The calculated value of ΔH_f° at 298.15 K is −245.23 kJ mol^{-1}, a value that differs from the experimental value by less than 1 kcal mol^{-1}.

7.3 Atomic Charges

The electrostatic potential for the H_2O molecule mapped with rainbow colors over a total density surface with isovalue equal to 0.00040 (Fig. 5.4) shows that the hydrogen atoms of the water molecule are strongly electrophilic and the oxygen atom is strongly nucleophilic.

The electrostatic potential of a molecule can be used to determine its distribution of **atomic charges**. To this end, the molecular electrostatic potential ϕ and the electrostatic potential η that results from a set of point charges q_α located on the molecule's nuclei are evaluated on the nodes i of a grid. For a neutral molecule like H_2O the sum of all the atomic charges is zero, and so the charge centered on the oxygen nucleus is −2 times the charge on each of the hydrogen atoms, $q_O = -2\ q_H$. Hence, in atomic units, η is given by

$$\eta_i = \sum_{\alpha} \frac{q_\alpha}{r_{\alpha i}} = -\frac{2q_H}{r_{Oi}} + \frac{q_H}{r_{H_1 i}} + \frac{q_H}{r_{H_2 i}} \tag{7.1}$$

where i represents a grid node. By the least squares optimization procedure, the atomic charge q_H can be determined by minimizing

$$\Theta(q_H, r_{Oi}, r_{H_1 i}, r_{H_2 i}) = \sum_i (\eta_i - \phi_i)^2 \tag{7.2}$$

with respect to q_H, that is, by solving

$$\frac{\partial \Theta}{\partial q_H} = 0 \tag{7.3}$$

The following table shows the molecular electrostatic potential (MEP) atomic charges and the Mulliken atomic charges for H_2O obtained from B3LYP calculations with different basis sets. It is shown that the introduction of polarized functions on oxygen and hydrogen atoms significantly improves the results, as can be expected. In addition, molecular electrostatic atomic charges obtained from large basis sets with diffuse functions give the same atomic charges. Note that Mulliken charges vary with the basis sets somewhat erratically.

	MEP		Mulliken	
Method/basis set	q_O	q_H	q_O	q_H
B3LYP/6-31G	−0.897	0.448	−0.718	0.359
B3LYP/6-31G(d,p)	−0.737	0.368	−0.610	0.305
B3LYP/6-311G(d,p)	−0.749	0.375	−0.474	0.237
B3LYP/cc-pVTZ	−0.703	0.352	−0.435	0.218
B3LYP/cc-pVQZ	−0.698	0.349	−0.490	0.245
B3LYP/aug-cc-pVTZ	−0.680	0.340	−0.354	0.177
B3LYP/aug-cc-pVQZ	−0.680	0.340	−0.587	0.294

7.4 Dipole Moment

7.4.1 Electric Multipoles

In this section, we deal with simple charge distributions and their electrostatic potentials. We begin by considering charge distributions that correspond to pure electric multipoles (charge, dipole, quadrupole, and octupole) and their corresponding numbers of poles expressed as powers of 2 (Fig. 7.3). Since a charge can be either positive or negative, the number of poles is given by a power of 2.

The contours for the electrostatic potential of a **charge** q, in the xy-plane and atomic units, are given by

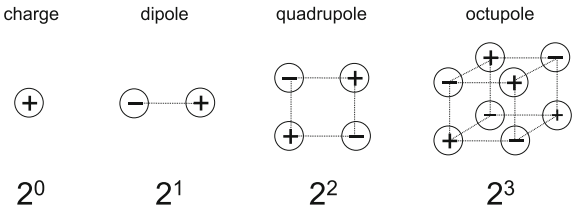


Fig. 7.3 Electric multipoles (schematic) with the corresponding number of poles expressed as a power of 2

$$\phi(x, y) = \frac{q}{r} \tag{7.4}$$

where r represents the distance between the point P with coordinates $(x, y, 0)$ and the charge located at the origin (0, 0, 0). These contours are concentric circles whose center is the point charge q (Fig. 7.4), since the electrostatic potential produced by a point charge depends only on the distance to the charge. All the radial directions emerging from the charge display the same variation of the electrostatic potential, that is, they are equivalent directions, and the space around the central charge is said to be **isotropic**.

The electrostatic potential of a **dipole** is

$$\phi(x, y) = q\left(\frac{1}{r_{1,0}} - \frac{1}{r_{-1,0}}\right) \tag{7.5}$$

where q represents the absolute value of the charges (the minus sign in the second member of this expression accounts for the sign of the corresponding negative charge), and $r_{-1,0}$ and $r_{1,0}$ represent the distances between the point P in the xy-plane and the negative and positive charges with coordinates $(-1, 0, 0)$ and $(1, 0, 0)$, respectively (Fig. 7.4; see *Mathematica* codes **M2** and **M3**; **E1**, **E2**, **E3**).

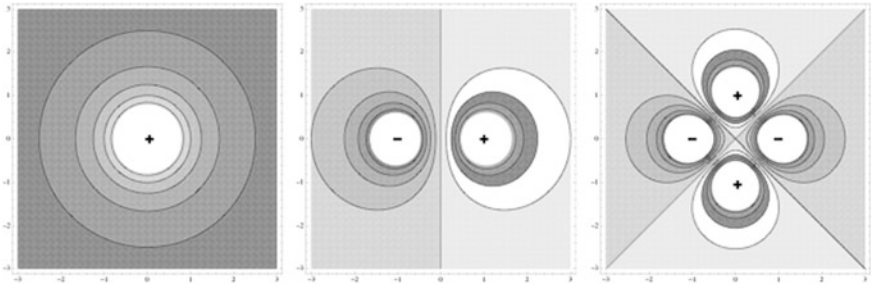


Fig. 7.4 Contours in the xy-plane for electrostatic potentials produced by one charge, one dipole, and one quadrupole. Figures obtained with Mathematica

The electrostatic potential of a **quadrupole** is given by (Fig. 7.4).

$$\phi(x, y) = q\left(\frac{1}{r_{0,1}} + \frac{1}{r_{0,-1}} - \frac{1}{r_{-1,0}} - \frac{1}{r_{1,0}}\right) \tag{7.6}$$

7.4.2 Point Dipole

As mentioned earlier, two electric charges $+q$ and $-q$ separated by a small distance d form an electric dipole. We now represent the distances between the positive and the negative charges, on the one hand, and the point P in the xy-plane, on the other, by r_+ and r_- (Fig. 7.5), and write the electrostatic potential as

$$\phi(x, y) = q\left(\frac{1}{r_+} - \frac{1}{r_-}\right) = q\frac{r_- - r_+}{r_+ r_-} \tag{7.7}$$

Substitution of

$$r_+ = \left[\left(x - \frac{\delta}{2}\right)^2 + y^2\right]^{1/2} \quad r_- = \left[\left(x + \frac{\delta}{2}\right)^2 + y^2\right]^{1/2} \tag{7.8}$$

in (7.7) followed by Taylor expansion of the resulting function of δ about the point $\delta = 0$ gives

$$\phi(\delta) = \phi(0) + \delta\phi'(0) + \frac{\delta^2}{2!}\phi''(0) + \frac{\delta^3}{3!}\phi'''(0) + \cdots \tag{7.9}$$

where

$$\phi(0) = 0 \quad \phi'(0) = \frac{qx}{r^3} \quad \phi''(0) = 0 \quad \phi'''(0) = \frac{qx(15x^2 - 9r^2)}{4r^7} \tag{7.10}$$

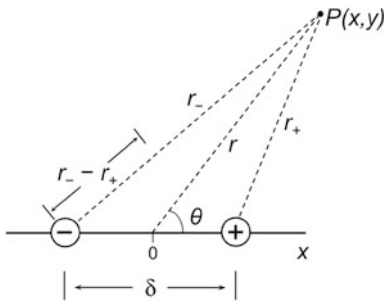


Fig. 7.5 Geometric variables for a dipole

In turn, substitution of (7.10) in (7.9) leads to

$$\phi(\delta) = \frac{q\,\delta\,x}{r^3} + \frac{q\,\delta^3\,x(5x^2 - 3r^2)}{8r^7} + O(\delta^5) \tag{7.11}$$

Since the exchange of positive and negative charges changes the sign of the electrostatic potential, this is an odd function of δ [$\phi(\delta) = -\phi(-\delta)$], and consequently, the even powers of δ in (7.9) are zero. If δ represents an arbitrarily small separation between the charges of the dipole, then we can approximate $\phi(\delta)$ to the first term of (7.11) (the terms in δ^3 and O(δ^5) become negligible compared with the δ term) and write

$$\mu = q\,\delta \tag{7.12}$$

where μ is the magnitude of the **point dipole** vector. Note that $x = r\cos\theta$ (**E4**).

7.4.3 Electric Field Streamlines

From the definition of electrostatic potential, (see 5.10), one can write

$$V = q_t\phi \tag{7.13}$$

where V is the potential energy, q_t is the test charge, and ϕ is the electrostatic potential. The x and y components of the electric force are given by

$$F_x = \frac{\partial V}{\partial x} = -q_t\frac{\partial \phi}{\partial x} \quad F_y = -\frac{\partial V}{\partial y} = -q_t\frac{\partial \phi}{\partial y} \tag{7.14}$$

and the x and y components of the electric field are

$$\mathcal{E}_x = \frac{F_x}{q_t} \quad \mathcal{E}_y = \frac{F_y}{q_t} \tag{7.15}$$

Substitution of (7.15) in (7.14) leads to

$$\mathcal{E}_x = -\frac{\partial \phi}{\partial x} \quad \mathcal{E}_y = -\frac{\partial \phi}{\partial y} \tag{7.16}$$

Note that

$$\frac{\partial \phi}{\partial x}\hat{\boldsymbol{i}} + \frac{\partial \phi}{\partial y}\hat{\boldsymbol{j}} = \nabla\phi \tag{7.17}$$

is the **gradient** vector of the electrostatic potential in the xy-plane, and ∇ represents the del, or nabla, operator indicative of the vector differential operator.

Consider now a contour in the xy-plane described by the x and y variables as functions of a parameter t (an external variable). If t represents the variable time, as it usually does in physical problems, then

$$\boldsymbol{v} = \left(\frac{\partial x}{\partial t}\right)_{t_0} \hat{\boldsymbol{i}} + \left(\frac{\partial y}{\partial t}\right)_{t_0} \hat{\boldsymbol{j}} \tag{7.18}$$

is the velocity vector tangent to the contour at (x_0, y_0) (this point corresponds to time t_0). The dot product of the gradient of the electrostatic potential and the velocity vector is given by

$$\nabla\phi.\boldsymbol{v} = \left(\frac{\partial \phi}{\partial x}\right)_{x_0,y_0}\left(\frac{\partial x}{\partial t}\right)_{t_0} + \left(\frac{\partial \phi}{\partial y}\right)_{x_0,y_0}\left(\frac{\partial y}{\partial t}\right)_{t_0} = \left(\frac{d\phi}{dt}\right)_{t_0} = 0 \tag{7.19}$$

This dot product is zero, because the electrostatic potential does not vary on a contour curve by the definition of contour. Hence, at (x_0, y_0), the gradient of the electrostatic potential is perpendicular to the velocity vector, which in turn is tangent to the contour at that point. The gradient vector points in the direction of the greatest rate of increase of the electrostatic potential. From (7.16) and (7.17), one concludes that the electric field vector in the xy-plane is

$$\mathcal{E}_x \hat{\boldsymbol{i}} + \mathcal{E}_y \hat{\boldsymbol{j}} = -\nabla\phi \tag{7.20}$$

and so, at a defined point of the xy-plane, the electric field vector points in the direction of the greatest decrease of the electrostatic potential. Figure 7.6 shows a plot of the electrostatic potential contours and the **electric field streamlines** for a dipole (Fig. 7.6, *left*) and a point dipole (Fig. 7.6, *right*; see *Mathematica* code **M4**). It can be seen that the streamlines move away from the positive charge (the **source**) and converge to the negative charge (the **sink**), since these are the directions of the greatest decrease of the electrostatic potential. In the case of the **point dipole**, source and sink coincide (the streamlines move away and converge to the point dipole).

7.4.4 H_2O Dipole and Quadrupole

We now consider the atomic charges for the water molecule derived from the electrostatic potential obtained from a *Gaussian* 09 B3LYP/aug-cc-pVTZ calculation (see Sect. 7.3). Since the electrostatic potential is an additive function, it can be equated to a sum of electrostatic potentials associated with nonnull electric multipoles and represented by the contours in the molecular plane. Figure 7.7 shows the contours of the total electrostatic potential of H_2O as an approximate sum of the contours for the electrostatic potential of the dipole term associated with the vector

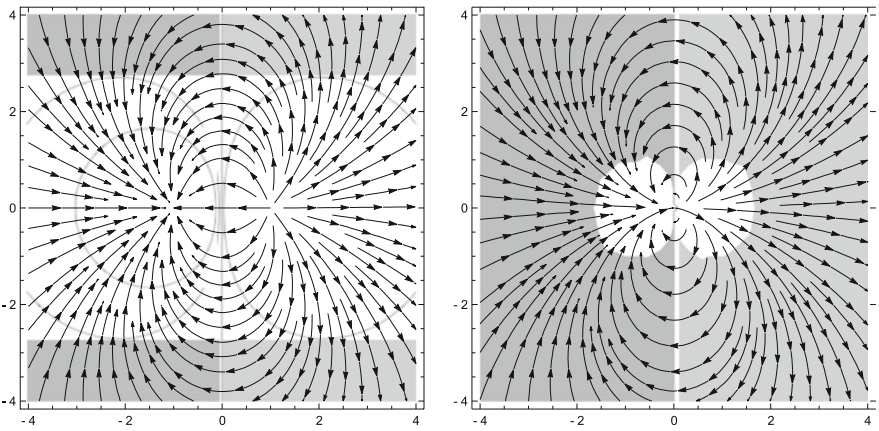


Fig. 7.6 Contours and streamlines show the direction of the electric field, for a dipole (*left*) and a point dipole (*right*). Figures obtained with *Mathematica*

sum of the O–H bond dipoles and the electrostatic potential for a quadrupole term (see **E5**). These results show that besides the dipole, the H_2O molecule has also a significant contribution from the quadrupole.

7.5 Molecular Orbitals

In Hartree–Fock and DFT calculations, the wave function for a closed shell molecule is a Slater determinant of one-electron wave functions (spin-orbitals) that are eigenfunctions of effective one-electron Hamiltonians called Fock and Kohm–

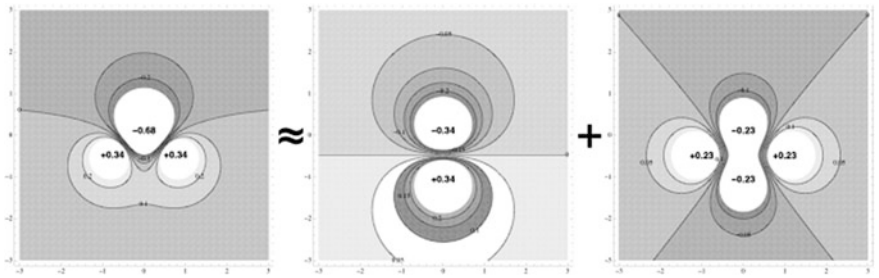


Fig. 7.7 Contours for the electrostatic potential of the H_2O molecule derived from the distribution of atomic charges (*Gaussian*09 B3LYP/aug-cc-pVTZ calculation) as an approximate sum of the contours for a dipole and a quadrupole. Figures obtained with *Mathematica*

Sham operators (**E6**). For a molecule, these one-electron wave functions are called **molecular orbitals** and are expressed as linear combinations of basis functions centered at each atom of the molecule.

C_{2v} character table

	E	C_2	σ_{xz}	σ_{yz}		
A_1	1	1	1	1	z	x^2, y^2, z^2
A_2	1	1	−1	−1	R_z	xy
B_1	1	−1	1	−1	x, R_y	xz
B_2	1	−1	−1	1	y, R_x	yz

(http://symmetry.jacobs-university.de)

The H_2O molecule belongs to the C_{2v} symmetry group, whose symmetry elements are shown in Fig. 7.8. The symmetry species of this group are presented in the corresponding character table (see above). Upon application of the C_{2v} symmetry operations, the H_2O molecular orbitals transform as the group symmetry species. For this reason, the molecular orbitals are labeled by the corresponding symmetry species. Being one-electron functions, they are labeled by lowercase letters.

The water molecule has ten electrons that are allocated to the five molecular orbitals with lower energy (each molecular orbital is occupied by two electrons with antiparallel spins). The five doubly occupied molecular orbitals of H_2O are shown in Fig. 7.9 by surfaces of defined isovalues, with the corresponding energies in eV,

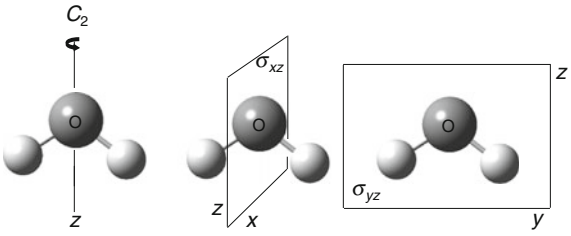


Fig. 7.8 Symmetry elements for the H_2O molecule. Molecular representations obtained with *GaussView*

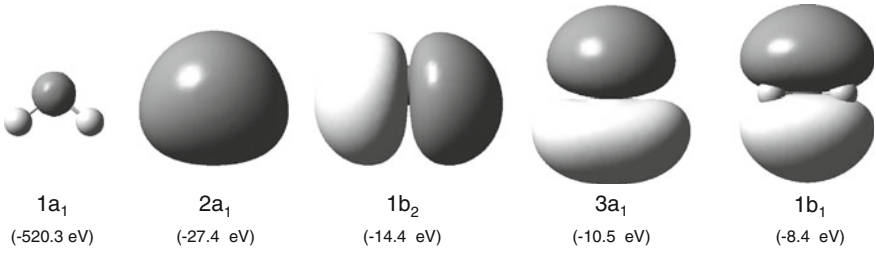


Fig. 7.9 Molecular orbitals for the H_2O molecule represented by surfaces whose isovalue is equal to 0.02000 (*Gaussian* 09 B3LYP/cc-pVTZ calculation)

in parentheses, and a_1, b_1, b_2 name the molecular orbitals by their symmetry types. The orbitals that have a common symmetry label are numbered by increasing energy (for instance, $1a_1$, $2a_1$, $3a_1$). The occupied molecular orbital of lowest energy ($1a_1$) is essentially restricted to the 1s oxygen atom core electrons, which are not significantly involved in the O–H bonds.

7.5.1 Natural Bond Orbitals

Figure 7.10 shows the natural bond orbitals BD_1 and BD_2 and the natural atomic orbitals LP_1 and LP_2 for the H_2O molecule. These natural orbitals have the following approximate expressions:

$$\begin{aligned} LP_2 &\approx p_O \\ LP_1 &\approx sp_O^{0.86} \\ BD_2 &\approx 0.85\, sp^{3.33}_{O\rightarrow H_2} + 0.52\, s_{H_2} \\ BD_1 &\approx 0.85\, sp^{3.33}_{O\rightarrow H_1} + 0.52\, s_{H_1} \end{aligned} \tag{7.21}$$

The natural bond orbitals BD_1 and BD_2 are approximate localized in the OH bonds and are equivalent orbitals by molecular symmetry, since the OH bonds of H_2O are equivalent bonds. In turn, LP_1 and LP_2 are approximate localized in the oxygen atom (lone pairs or natural atomic orbitals).

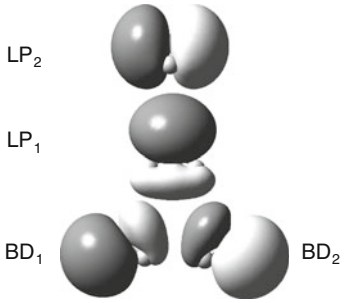


Fig. 7.10 Surfaces of isovalue equal to 0.0200 represent the natural bond orbitals BD_1 and BD_2 and the oxygen atom lone pairs LP_1 and LP_2 (*Gaussian* 09 B3LYP/cc-pVTZ calculation)

7.6 Molecular Vibrations

The H_2O molecule has $N = 3$ atoms and consequently $3N - 6 = 3$ vibrational coordinates. Two of these are O–H stretchings, the third being the H–O–H angle bending. According to the H_2O symmetry (C_{2v} symmetry group), the equivalent O–H oscillators cannot oscillate separately: they oscillate either in phase (**symmetric stretching**) or out of phase (**antisymmetric stretching**). In the symmetric stretching, both H atoms move apart from the oxygen atom or get closer to it synchronously, that is, in phase. Apart from a normalization factor, the symmetric stretching coordinate is the sum of the OH stretching coordinates and transforms like the A_1 symmetry species. In turn, in the antisymmetric stretching, one of the H atoms moves apart from the oxygen atom and the other H atom gets closer to it, that is, they move out of phase. Both pass by their extreme positions at the same time. Apart from a normalization factor, the antisymmetric stretching coordinate is the difference of the OH stretching coordinates and transforms like the B_2 symmetry species. The H–O–H **bending** is a totally symmetric vibration, that is, it transforms as the symmetry species A_1.

The vibrational modes of a molecule that transform like the species of its symmetry group are called **normal vibrational modes**. The antisymmetric stretching, the symmetric stretching, and the bending are the normal vibrational modes of the H_2O molecule (Fig. 7.11) whose scaled wavenumbers (the scaling factor for the calculation B3LYP/cc-pVTZ is equal to 0.965; see CCCBDB) are 3767, 3670 and 1581 cm^{-1}, respectively. These frequencies are less than 1% away from the experimental values (3756, 3657, 1595 cm^{-1}; see CCCBDB).

Displacement vectors and scaled wavenumbers for the calculated normal vibrational modes of H_2O

ν/cm^{-1}		3767			3670			1581	
	x	*y*	*z*	*x*	*y*	*z*	*x*	*y*	*z*
O	0.00	0.07	0.00	0.00	0.00	0.05	0.00	0.00	0.07
H	0.00	**−0.56**	**0.43**	0.00	**0.59**	**−0.39**	0.00	**−0.43**	**−0.56**
H	0.00	**−0.56**	**−0.43**	0.00	**−0.59**	**−0.39**	0.00	**0.43**	**−0.56**

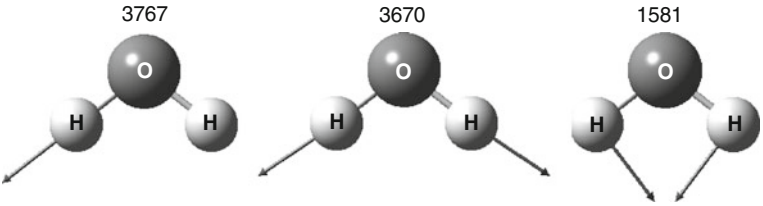


Fig. 7.11 Normal vibrational modes of H_2O with the wavenumbers in cm^{-1} after scaling by the factor 0.965 (CCCBDB) and the corresponding displacement vectors (*Gaussian* 09 B3LYP/cc-pVTZ calculation). In the normal mode at 3767 cm^{-1}, the displacement vector for the OH oscillator on the right has its origin on the H atom and is directed toward the O atom, thus being hidden by the segment that represents the OH bond

The calculated displacement vectors are shown in the above table. Note that the *x*-axis is perpendicular to the molecular plane (the *yz*-plane) (see Fig. 7.8). Since the displacement vectors lie in the molecular plane, all the *x*-components of the displacement vectors are zero. The numbers highlighted in bold correspond to the largest displacements of the hydrogen atoms, and the small displacements of the oxygen atom move this atom so that the center of mass of the molecule stays unshifted (a center of mass motion is a translation, not a vibration).

Displacement vectors and scaled wavenumbers for the calculated normal modes of vibration of DOH

ν/cm^{-1}		3720			2701			1386	
	x	*y*	*z*	*x*	*y*	*z*	*x*	*y*	*z*
O	0.00	−0.05	0.03	0.00	0.09	−0.07	0.00	0.00	0.09
H	0.00	**0.81**	**−0.59**	0.00	0.11	−0.01	0.00	**−0.55**	**−0.70**
D	0.00	0.01	0.03	0.00	**−0.80**	**−0.58**	0.00	**0.28**	**−0.35**

Protium (^{1}H) and deuterium (D = ^{2}H) have the same atomic number $Z = 1$, and so both are hydrogen atoms. Within the Born–Oppenheimer approximation, the H_2O and DOH molecules have the same molecular orbitals. However, the vibrational frequencies depend on the reduced masses of the oscillators, and a protium by deuterium substitution breaks the equivalence of the oxygen–hydrogen oscillators. The reduced masses of the OH and OD oscillators in unified mass units are approximately equal to $16 \times 1/(16+1) = 16/17$ and $16 \times 2/(16+2) = 32/18$, respectively. Hence,

$$\frac{\omega_{OH}}{\omega_{OD}} \approx \sqrt{\frac{\mu_{OD}}{\mu_{OH}}} \approx \sqrt{2} \tag{7.22}$$

(**E7**).

The calculated scaled wavenumbers and displacement vectors for the normal modes of DHO are shown in Fig. 7.12. The normal modes at 3720 and 2701 cm^{-1} have predominant contributions of the OH and OD stretching vibrations, respectively (see table of atomic displacements). The large variation in the reduced mass

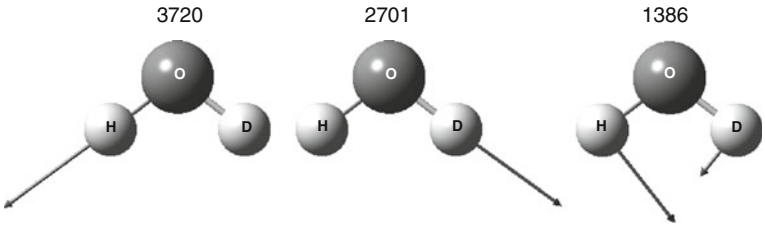


Fig. 7.12 Normal vibrational modes of DHO with the wavenumbers in cm^{-1} after scaling with the factor 0.965 (CCCBDB), and the corresponding displacement vectors (*Gaussian* 09 B3LYP/cc-pVTZ calculation)

of the oxygen–hydrogen oscillator that resulted from the protium by deuterium substitution has the effect of localizing the OH and OD stretching vibrations. Hence, the vibrational mode at 3720 cm^{-1} is essentially the OH stretching, whereas the mode at 2701 cm^{-1} has a predominant contribution from the OD stretching (note the wavenumbers' shift of more than 100 cm^{-1} with respect to the OH stretching) (**E8**, **E9**).

7.7 Intermolecular Interactions

The existence of intermolecular interactions can be easily inferred from simple experimental observation: gases would not liquefy and liquids would not solidify if there were no attractive long-range intermolecular interactions. Repulsive interactions occur at short range and prevent solids and liquids from being compressed into much smaller volumes. Hence, long-range and short-range interactions can be inferred from the simple existence of liquids and solids and their reduced compressibility.

Intermolecular interactions can be classified as electrostatic, induction, and dispersion, according to the physical phenomena that produce their long-range behavior, which leads to energies that vary with powers of inverse distance. In turn, short-range interactions are mostly due to repulsion and exchange between electrons of different molecules, and they decrease exponentially with distance. Long-range and short-range interactions can be illustrated by the potential energy curves for the interaction between two closed-shell atoms or monatomic ions, since the corresponding potential energy curves provide useful insight into the fundamental features of the interactions.

7.7.1 Electrostatic Interaction

Sodium and fluoride ions have the same number of electrons (10) and the same closed-shell electron configuration, $1s^2 2s^2 2p^6$. The potential energy curve for $Na^+ \ldots F^-$ shows a deep minimum at approximately 2.0 Å that results from a balance between repulsive and attractive interactions (Fig. 7.13). For interatomic distances smaller than that of the minimum, the dominant interaction has its origin in the repulsion and exchange between electrons of different atoms (exchange of electrons of different atoms requires overlap between the electron densities of both atoms) and corresponds to a curve that decreases with increasing internuclear distance. In turn, the leading long-range interaction between the positive and negative charges of Na^+ and F^- is **electrostatic**, mostly of the charge-to-charge type. This interaction is approximately described by a $-1/R$ curve that represents the attractive wing of the potential well (Fig. 7.13) and tends asymptotically to zero, thus meaning that the additive constant in the $-1/R$ fitting is set to zero energy. At $R = 6.0$ Å, the potential energy curve is well below zero, and the charge-to-charge electrostatic

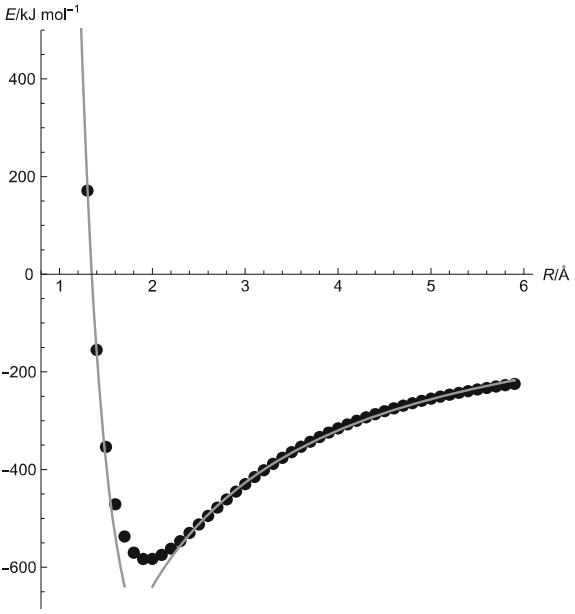


Fig. 7.13 Potential energy curve for $Na^+ \ldots F^-$ (*Gaussian* 09 B3LYP/ aug-cc-pVTZ calculation). Curves are separately fitted to short-range repulsive and long-range attractive wings of the potential energy curve. The long-range fitting curve corresponds to the $1/R$ distance dependence. The zero energy was chosen so that the fitting curve tends to zero as $R \to \infty$. Figure obtained with *Mathematica*

interaction is still strong, since electrostatic interactions are relatively strong interactions.

Electrostatic interactions involve permanent electric moments of the interacting molecules, and generally speaking can be of the charge-to-charge, charge-to-dipole, dipole-to-dipole, dipole-to-quadrupole, … types. Usually, the leading electrostatic interaction comes from the first significant electric moment in each of the interacting molecules. The interaction between permanent 2^n-pole and 2^m-pole electric moments varies as $1/R^{(n+m+1)}$. For a charge–charge interaction, $n = 0$ and $m = 0$, and the distance-dependence is given by $1/R$. A charge–dipole interaction corresponds to n=0 and m=1 with the resulting electrostatic interaction varying with distance as $1/R^2$. In turn, a dipole–dipole interaction corresponds to $n = 1$ and $m = 1$, that is, the interaction varies with distance as $1/R^3$. Electrostatic interactions can be attractive or repulsive and involve permanent electric moments of the interacting molecules, thus being pairwise additive. For an assembly of molecules A, B, and C, pairwise addition means that the total electrostatic energy is given by the sum of the electrostatic interactions associated with the permanent electric moments of the AB, BC, and CA pairs.

Since the electron density of atoms or molecules is polarizable, a permanent electric moment in one molecule gives rise to an induced electric moment into the other molecule. Contributions like these are included in the induction interaction, as we now consider.

7.7.2 Induction

The sodium ion is isoelectronic with a neon atom, and both have the same number of electrons (10) and the same electronic configuration, $1s^2 2s^2 2p^6$. *Gaussian* calculations for the $Na^+ \ldots Ne$ system were carried out using the Moller–Plesset perturbation method to fourth order of perturbation (MP4) with the aug-cc-pVTZ basis set [MP4(sdq)/aug-cc-pVTZ calculation]. The potential energy curve for $Na^+ \ldots Ne$ (Fig. 7.14) shows an energy minimum at 2.55 Å whose depth is more than one order of magnitude smaller than the minimum observed for $Na^+ \ldots F^-$. The charge of Na^+ induces a dipole in the neon atom with the negative end closer to the sodium ion. The resulting charge-induced dipole attractive interaction is called **induction**.

In general, the interaction between a permanent 2^n-pole and an induced 2^m-pole varies as $1/R^{2(n+m+1)}$. For a charge-induced dipole interaction, n=0 and m=1, the distance dependence is given by $1/R^4$. Induction interactions are clearly nonadditive, since the multipole moments induced by a third molecule interact, being subject to mutual change. The weak solubility of oxygen in water can be approximately described by the interaction between the dipole moment of the water molecule and the induced dipole moment in the diatomic oxygen molecule.

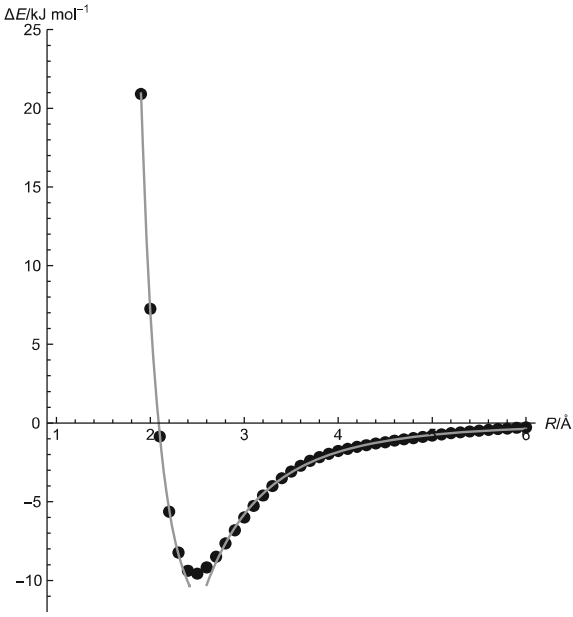


Fig. 7.14 Potential energy curve for $Na^+ \ldots Ne$ [*Gaussian* 09 MP4(sdq)/aug-ccpVTZ calculation]. Curves are separately fitted to the short-range repulsive and long-range attractive wings of the potential energy curve. The long-range fitting curve corresponds to the $1/R^4$ dependence. The zero energy was chosen so that the $1/R^4$ fitting curve tends to zero as $R \rightarrow \infty$. Figure obtained with *Mathematica*

7.7.3 Dispersion

Helium and neon boil at 4.22 K and 27.07 K, with these boiling temperatures corresponding to *RT* values approximately equal to 0.04 and 0.2 kJ mol^{-1}. Since these energies are well below 1 kcal mol^{-1} (≈4.2 kJ mol^{-1}), the calculation of the interaction energy between two rare gas atoms cannot be obtained from the difference between the total energy of the system of the interacting atoms ($E_{A\ldots A}$) and the sum of the individual energies of each atom ($2E_A$), since the errors of these large numbers are far greater than the resulting energy difference. The interaction between two rare gas atoms corresponds to a tiny perturbation, and so should be evaluated as such by a perturbation method. To this end, *Gaussian* 09 calculations for the He…He and Ne…Ne systems were carried out using the Moller–Plesset perturbation method to fourth order of perturbation (MP4) with the aug-cc-pVTZ basis set [MP4(sdq)/aug-cc-pVTZ calculation] (Aziz and Slaman 1991; Woon 1991). The potential energy curves for He…He and Ne…Ne systems present very shallow minima (approximately 0.07 and 0.2 kJ mol^{-1}) at interatomic distances around 3.05 and 3.15 Å, respectively, with the minimum for Ne…Ne being about nine times deeper than for He…He (see insets in the potential energy curves of Fig. 7.15).

Isolated atoms are isotropic systems with respect to their nuclei, and all permanent electric moments of neutral atoms are null. Therefore, the observed minima

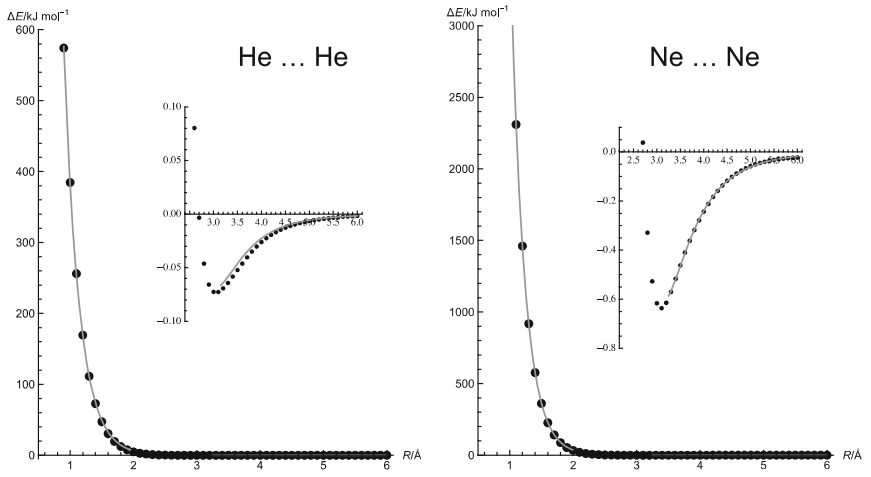


Fig. 7.15 Potential energy curves for He…He and Ne…Ne as functions of the internuclear distance [*Gaussian* 09 MP4(sdq)/aug-ccpVTZ calculation]. Curves are fitted to the short-range repulsive wings of the potential energy curve. Details of the minima and the curves fitted to the attractive wings of the very shallow potential wells are shown in the insets. Energy zeros were chosen so that the long-range fitting curves tend to zero as $R \rightarrow \infty$. Figures obtained with *Mathematica*

in the potential energy curves of He…He and Ne…Ne systems cannot derive from interactions involving permanent electric moments. What is the origin of the observed minima? Why is the minimum of Ne…Ne deeper than that of He…He?

Electrons try to avoid each other due to instantaneous repulsions. While the instantaneous asymmetry in the electron distribution of a single atom averages to zero over time, the asymmetry in the electron distribution of one atom produced by correlation between electrons of different atoms can be described in each atom by successive induced electric moments (induced dipole, induced quadrupole, …) and depends on the internuclear distance. Figure 7.16 schematically illustrates how an instantaneous dipole in the electron distribution of one atom induces a dipole in the electron distribution of the other atom. The resulting induced dipole-induced dipole interaction is attractive, since induced dipoles always reorient to produce attraction. In general, the interaction between induced 2^n- and 2^m-poles varies as $1/R^{2(n+m+1)}$. An induced dipole-induced dipole interaction ($n = 1$ and $m = 1$) corresponds to $1/R^6$ dependence. The correlation between electrons of different atoms can also result from the contributions of induced electric moments of order greater than 1, for instance, induced dipole-induced quadrupole ($1/R^8$) and induced quadrupole-induced quadrupole ($1/R^{10}$) interactions. The overall resulting attraction is called **dispersion**. Due to correlation between electrons of different atoms, neon atoms lead to a larger well depth than helium atoms, since the dispersion energy increases with the number of electrons of the interacting atoms. An example of this is the fact that chlorine is a gas, bromine a liquid, and iodine a solid.

Since the electron densities of interacting atoms or molecules are polarizable, dispersion always occurs, that is, it is found in every molecular interaction. Its $1/R^6$ distance-dependence makes it approximately additive, despite being due to an interaction between induced electric moments. For small or medium-size molecules, dispersion is a weak interaction, but it becomes significant and sometimes even decisive when the remaining interactions are null or negligible.

Dispersion interactions between alkane molecules contribute appreciably to determining the physical state of a hydrocarbon. In fact, the magnitude of the dispersion interactions increases with the number of electrons in the atoms located on the molecular surface (for alkane molecules, these are hydrogen atoms). This means that the magnitude of the dispersion interaction increases with the area of the molecular surface and the number of hydrogen atoms on it. For example, 2,2-dimethyloctane (a branched decane) has a normal boiling point at 155 °C,

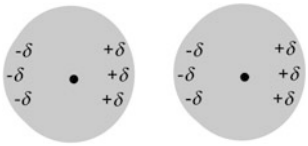


Fig. 7.16 Correlation between electrons of different atoms leads to attraction involving an induced dipole-induced dipole interaction (schematic)

whereas the boiling point of *n*-decane is 174 °C (the linear alkane has a larger surface area than the branched alkane).

The boiling points of linear alkanes are plotted in Fig. 7.17 as a function of the number of hydrogen atoms. Since the general formula for an alkane is C_nH_{2n+2}, an increase of one carbon atom corresponds to an increase of two hydrogen atoms, and linear alkanes that have four or fewer carbon atoms are gases at ambient temperature. Long hydrocarbon chain molecules with a large number of hydrogen atoms in the interfacial area between molecules have appreciable dispersion energies. This can be inferred from their high boiling points: at ambient temperature and normal pressure, linear alkanes with more than seven carbon atoms are liquids, and linear alkanes with more than 17 carbon atoms are solids.

Electrostatic, induction, dispersion, and exchange-repulsion interactions are *fundamental* interactions. We now consider a *specific* interaction in which the hydrogen atom exerts a major role: the hydrogen bond.

7.8 Hydrogen Bond

Consider one hydrogen fluoride molecule interacting with one water molecule. The potential energy curve as a function of the H…O distance (Fig. 7.18) presents an energy minimum in which the hydrogen atom covalently bonded to the fluorine atom interacts with the oxygen atom of the water molecule. The relaxed scan of the H…O distance leads to an energy minimum whose depth is approximately 37 kJ mol^{-1} (*Gaussian* 09 B3LYP/aug-cc-pVTZ calculation). Acting like a bridge between two electronegative atoms, fluorine and oxygen, the hydrogen atom establishes a weak directional interaction with one of the lone pairs in the oxygen atom. At the minimum, the F–H bond length (0.945 Å) is slightly longer than the bond length of an isolated hydrogen fluoride molecule (0.924 Å), and the H…O

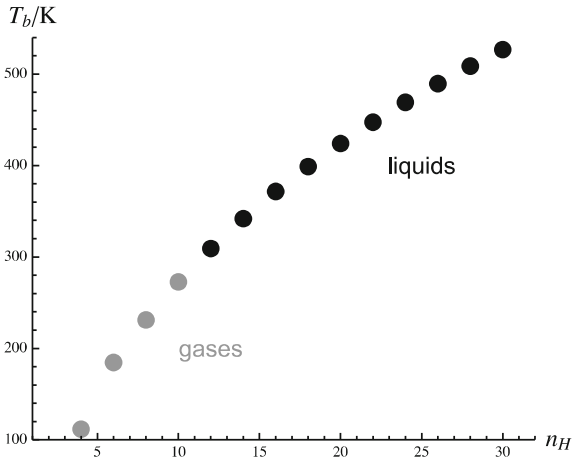


Fig. 7.17 Normal boiling points of linear alkanes versus the number of hydrogen atoms. Figure obtained with *Mathematica*

internuclear distance (1.692 Å) almost doubles the F–H bond length. The F–H…O interaction is a **hydrogen bond** and the FH…OH_2 complex can be called a hydrogen bond **heterodimer**, with the HF and H_2O molecules having distinct roles: HF provides the hydrogen atom (it is called the **donor**), whereas H_2O provides the lone pair in the oxygen atom (it is called the **acceptor**).

The hydrogen chloride–water system provides another example of a hydrogen bond heterodimer, ClH…OH_2. This time, the relaxed scan of the H…O distance leads to a minimum, whose calculated depth is approximately 20 kJ mol^{-1} (*Gaussian* 09 B3LYP/aug-cc-pVTZ calculation).

Isodensity surfaces mapped with electrostatic potentials can be used to illustrate the hydrogen bond formation in the FH…OH_2 heterodimer (Fig. 7.19). The electrostatic attractive interaction between the electrophilic region of the hydrogen atom in the hydrogen fluoride molecule and the nucleophilic region of the oxygen atom in the water molecule leads to the formation of the hydrogen bond heterodimer. The complementary functions of the donor and acceptor molecules are well illustrated by this electrostatic potential color map.

Note that each hydrogen bond heterodimer corresponds to a single closed isodensity surface, that is, the hydrogen fluoride and the water moieties do not have distinct isodensity surfaces. This means that the H…O interaction has some **partial covalent character**, which leads to the directional properties of the hydrogen bond. The fact that the expression "hydrogen bond" proposed by M.L. Huggins in 1919 prevailed over "hydrogen bridge" is probably due to the recognition of the covalent character of the hydrogen bond. In addition, the heterodimer exhibits stronger and larger electrophilic regions in the hydrogen atoms of the water molecule than in the isolated water molecule and a stronger and larger nucleophilic region in the fluoride atom of the heterodimer than in the corresponding isolated hydrogen fluoride molecule. These features point to the **polarization** of the FH…OH_2 heterodimer with respect to the individual molecules.

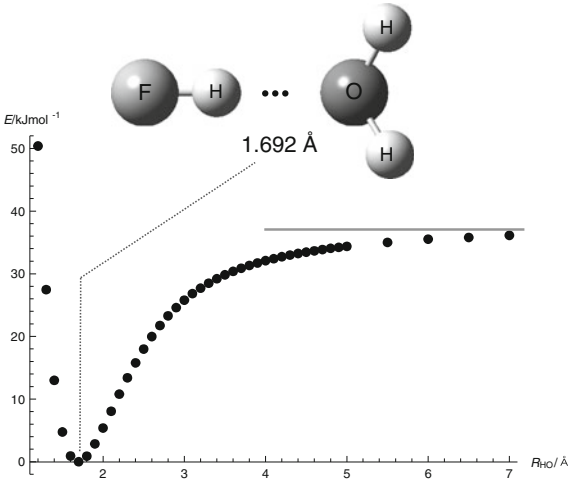


Fig. 7.18 Relaxed scan of the H…O distance in FH…OH_2 (*Gaussian* 09 B3LYP/aug-cc-pVTZ calculation). Figure obtained with *Mathematica*. Molecular representations obtained with *GaussView*

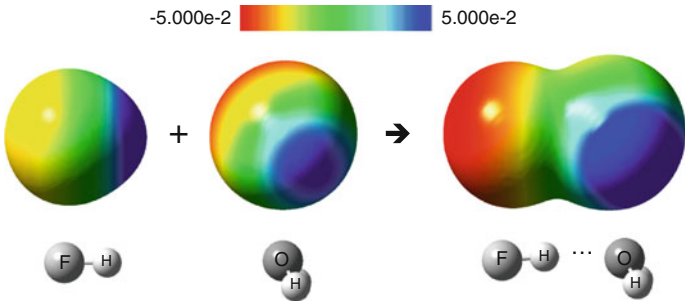


Fig. 7.19 Isodensity surfaces mapped with electrostatic potentials for HF and H_2O molecules, and the hydrogen bond heterodimer FH…OH_2 (*Gaussian* 09 B3LYP/aug-cc-pVTZ calculation)

Fig. 7.20 Atomic charges fitted to electrostatic potentials of HF, H_2O, and FH…OH_2 (*Gaussian* 09 B3LYP/aug-cc-pVTZ calculation)

Atomic charges fitted to electrostatic potentials, before and after hydrogen bond formation, are shown in Fig. 7.20. It can be seen that the hydrogen atoms of the water molecule are more positively charged in the hydrogen bond heterodimer than in the isolated water molecule, and the fluoride atom is more negatively charged in the hydrogen bond heterodimer than in the isolated hydrogen fluoride molecule. Thus, the distribution of atomic charges reveals polarization of the donor and acceptor molecules in the hydrogen bond heterodimer, as we already found by inspection of the electrostatic potential map. As a result of this polarization, the hydrogen fluoride and the water moieties are no longer neutral, due to the occurrence of charge transfer within the heterodimer, a fact that stresses the partial covalent character of the hydrogen bond. Only the sum of all the atomic charges in the whole hydrogen bond heterodimer is zero.

7.8.1 The Water Dimer

Consider the hydrogen bond water dimer, HOH…OH_2. Figure 7.21 shows the potential energy curve as a function of the H…O distance for the C_s configuration, where the plane of reflection contains the HOH…O nuclei. The relaxed scan of the H…O distance leads to an energy minimum, whose depth is approximately

19 kJ mol^{-1} (*Gaussian* 09 B3LYP/aug-cc-pVTZ calculation). In the O–H…O fragment, the O–H bond length (0.970 Å) is slightly longer than the O–H bond length of an isolated water molecule (0.960 Å), the H…O distance (1.953 Å) is slightly longer than double the O–H bond length, and the OO distance (2.917 Å) is slightly shorter than the sum of the O–H and H…O distances (2.923 Å), meaning that the O–H…O interaction is not exactly linear (the calculated value of the O–H…O angle is equal to 172.1°). These structural features follow trends similar to those found for the FH…OH_2 heterodimer.

The isodensity surfaces mapped with electrostatic potentials suggest that the electrostatic attraction between the electrophilic region on the HOH donor molecule and the nucleophilic region of the oxygen atom on the OH_2 acceptor molecule precedes the formation of the water dimer (Fig. 7.22). The electrophilic and nucleophilic regions are wider in the hydrogen bond dimer than in the isolated water molecule, pointing to the polarization of the water molecules in the water dimer. In addition, the individual water molecules that form the water dimer do not have distinct isodensity surfaces (Fig. 7.22), and the electrostatic potential in the internuclear H…O region is close to zero (note the green color), thus pointing to the partial covalent contribution of the hydrogen bond. In order to confirm this feature of the hydrogen bond, we now consider the intermolecular vibrations of the water dimer (Fig. 7.23).

The water dimer has $N = 6$ atoms, that is, $3N - 6 = 12$ vibrational modes. Six of these are intramolecular vibrations mostly localized in each water molecule, corresponding to the antisymmetric and symmetric stretchings and to the bending mode in each H_2O molecule of the dimer. The remaining six vibrational modes are intermolecular vibrations that comprise the out of plane and in plane bends, the hydrogen bond stretch, the twist, the wag, and the torsion vibrations (Fig. 7.23). By their own intermolecular nature, these vibrations extend to both water molecules, a

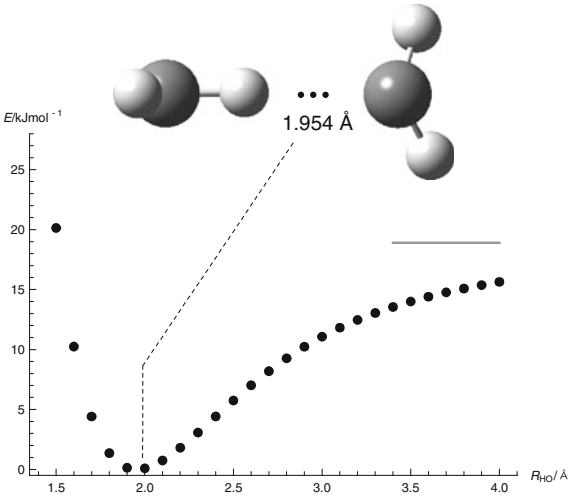


Fig. 7.21 Relaxed scan of the H…O distance in the hydrogen bond water dimer (*Gaussian* 09 B3LYP/aug-cc-pVTZ calculation). Figure obtained with *Mathematica*. Molecular representations obtained with *GaussView*

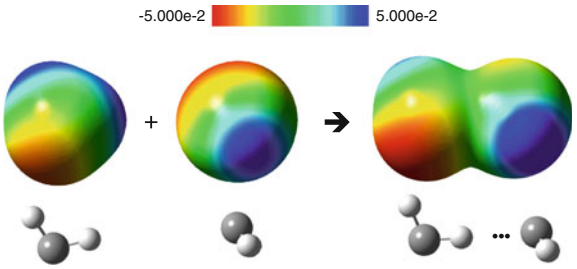
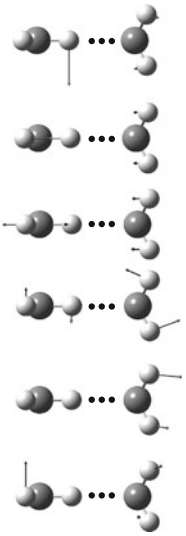


Fig. 7.22 Isodensity surfaces mapped with electrostatic potentials *before* and *after* hydrogen bond formation for the HOH…OH_2 dimer (*Gaussian* 09 B3LYP/aug-cc-pVTZ calculation)

Fig. 7.23 Intermolecular vibrational modes of the water dimer with calculated (*Gaussian* 09 B3LYP/aug-cc-pVTZ calculation) and experimental (CCCBDB) frequencies. Calculated frequencies involve scaling with the factor 0.968 (CCCBDB)



HOH … OH_2	ν_{calc} / cm^{-1}	ν_{exp} / cm^{-1}
out of plane bend (A")	607	523
in plane bend (A')	353	311
stretch (A')	185	143
twist (A")	156	108
wag (A')	155	103
torsion (A")	133	88

feature that confirms the covalent character of the hydrogen bond. In particular, the hydrogen bond stretching at 143 cm^{-1} (calculated frequency 185 cm^{-1}) is a consequence of the partial covalent character in the O…H interaction.

7.9 Ice I_h

In ice each water molecule has both donor and acceptor hydrogen bond functionalities, and it is involved in four hydrogen bonds directed to the vertices of a regular tetrahedron occupied by the nuclei of oxygen atoms of the neighboring water molecules. In fact, the water molecule at the center of this tetrahedral arrangement acts as a donor in two hydrogen bonds and as an acceptor in the other two hydrogen

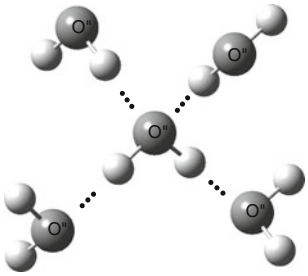


Fig. 7.24 Pentamer of water molecules in ice I_h. This structure is stable only as part of ice I_h

bonds (Fig. 7.24). Each water molecule at one vertex of the tetrahedral arrangement also takes the role of a central water molecule in a new tetrahedron not shown in Fig. 7.24. This tetrahedral pattern leads to the formation of hexagonal channels in **ice I_h** (I stands for 1 in roman numerals, and the subscript h stands for hexagonal) that result from the stacking of hexagonal cycles of hydrogen-bonded water molecules in the chair configuration (Fig. 7.25). At atmospheric pressure (1 atm = 101.325 kPa), the melting point of ice I_h is 273.15 K.

More than ten different forms of ice are known in distinct *p*-*T* regions of the phase diagram of water. **Ice polymorphism** (variety of hydrogen bonding architectures of water molecules in solid phase) results from the flexibility that structures based on O–H…O hydrogen bond networks have when subject to high pressures and very low temperatures.

7.9.1 Gas Hydrates

The diversity of solid structures of water is not restricted to ice forms. Under appropriate conditions of pressure and temperature, small gas molecules like CH_4, NH_3, and H_2S become trapped inside convex polyhedra formed by hydrogen-bonded water molecules. A convex polyhedron has no indentations or holes, that is, every linear segment connecting two points inside a convex polyhedron is entirely contained within it. The Swiss mathematician Leonhard Euler (1707–1783) discovered in 1751 that a convex polyhedron with V vertices, E edges, and F faces satisfies the equation $V - E + F = 2$. Oxygen atoms of hydrogen-bonded water molecules are at the vertices of the polyhedra whose edges correspond to O–H…O hydrogen bonds The sharing of faces by the convex polyhedra of hydrogen-bonded water molecules gives rise to solid forms of water that include small gas molecules and are called **gas hydrates**. Without the trapped molecules, the network of hydrogen-bonded water molecules would collapse into liquid water or a conventional ice form.

Large deposits of methane hydrates are found on the deep ocean floor of continental edges, at depths of hundreds of meters and pressures of hundreds of bars, near volcanic fissures from which gases like methane can burst. Methane from

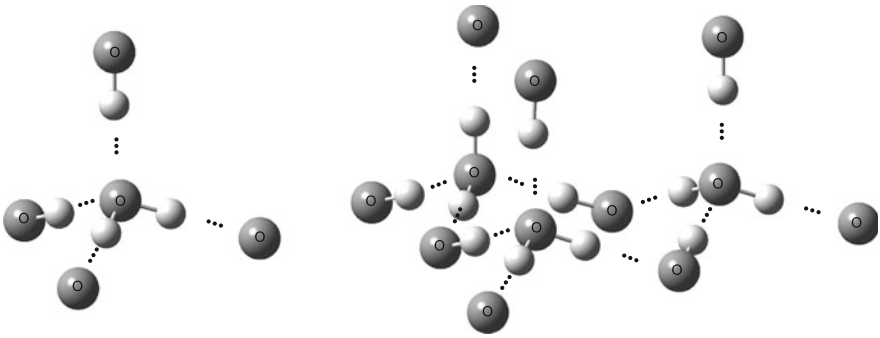


Fig. 7.25 *Left* Regular tetrahedron of hydrogen-bonded water molecules in ice seen as a pyramid. *Right* Ice fragment showing one hexagonal cycle of hydrogen-bonded water molecules in the chair configuration. For clarity, some hydrogen atoms covalently bonded to oxygen atoms are omitted

methane hydrate reservoirs doubles the amount of fossil fuel in the world (see Seuss et al. 1999). However, gas hydrates become unstable at temperatures above 0 °C and pressures below those found in the submarine deposits where they were formed.

7.10 Liquid Water

In the liquid state, the hydrogen-bonded network is dynamic, less ordered than in ice, and the hydrogen bonds around each water molecule are distorted with little directionality. In addition, the distribution of hydrogen bonds is strongly dependent on temperature. According to a simple model due to Luzar (see Luzar et al. 1983; Teixeira and Luzar 1999), it is assumed that in liquid water, no water molecule can form more than four hydrogen bonds with adjacent water molecules, that is, there are five groups of water molecules according to the number of hydrogen bonds each water molecule can form (4, 3, 2, 1, 0). In order to evaluate the population of each of these groups of water molecules at a defined temperature, we need to express the fraction of formed hydrogen bonds as a function of temperature. Neglecting hydrogen bond cooperativeness, each pair of oxygen atoms of adjacent water molecules can be in one of two states, one for a formed hydrogen bond, the other for a broken hydrogen bond. According to this model, the ratio of populations of these two states (broken over formed hydrogen bonds) is given by

$$\frac{(1-p_{hb})}{p_{hb}} = \alpha \exp\left(\frac{-\varepsilon}{RT}\right) \tag{7.23}$$

where p_{hb} is the probability of one hydrogen bond $1 - p_{hb}$ is the probability of a broken hydrogen bond, and α and ε are empirical parameters whose values $\alpha = 178$

and ε = 13.8 kJ mol^{-1} better describe the experimental results (**E10**). Rearranging (7.23), we obtain

$$p_{hb} = \frac{1}{1 + \alpha\ \exp[-\varepsilon/(RT)]} \quad 1 - p_{hb} = \frac{\alpha\ \exp[-\varepsilon/(RT)]}{1 + \alpha\ \exp[-\varepsilon/(RT)]} \tag{7.24}$$

The binomial distribution to obtain the fraction of water molecules in each of the five groups (k takes the values 4, 3, 2, 1, 0) is given by

$$\binom{4}{k} = \frac{4!}{k!(4-k)!} \quad 0 \le k \le 4 \tag{7.25}$$

For k = 4, 3, 2, 1, and 0, (7.25) takes the values 1, 4, 6, 4, 1, respectively. For example, there are six distinct ways [4!/ (2! 2!) = 6] for one water molecule to form two hydrogen bonds (Fig. 7.26).

Combining (7.24) and (7.25), we can write

$$\binom{4}{k} p_{hb}^{k}(1 - p_{hb})^{4-k} = \binom{4}{k} \frac{\{\alpha \exp[-\varepsilon/(RT)]\}^{4-k}}{\{1 + \alpha \exp[-\varepsilon/(RT)]\}^{4}} \quad 0 \le k \le 4 \tag{7.26}$$

Figure 7.27 shows the hydrogen bond distributions given by the Luzar model, at -30 °C (supercooled water), 10 °C, 50 °C and 90 °C, with maxima at k = 4, 3, 2, and 1 hydrogen bonds, respectively. The *Mathematica* code **M5** shows how the distribution of hydrogen bonds varies with temperature between 273 and 373 K, in steps of 1 K.

Fig. 7.26 The number of distinct ways for one water molecule to form two hydrogen bonds (*solid lines*) is given by 4!/(2! 2!) = 6

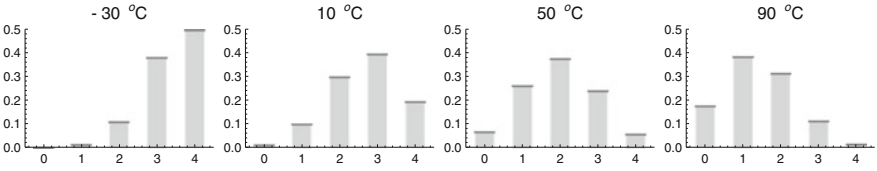


Fig. 7.27 Distribution of hydrogen bonds in water at defined temperatures, according to the Luzar model. This figure was obtained with *Mathematica*

The main limitations of the Luzar model are due to the fact that the binomial distribution is restricted to five groups of water molecules (k takes the discrete values 4, 3, 2, 1, and 0) and to the neglect of cooperativeness in hydrogen bond formation (p_{hb} does not depend on k). These limitations prevent us from extracting quantitative conclusions from the above results. Nevertheless, the plots show two important general trends, namely, the hydrogen bond distribution in liquid water is strongly dependent on temperature, and as temperature increases, the distribution maximum moves to lower numbers of hydrogen bonds.

7.11 Phase Diagram

The phase diagram of water shows ice–liquid, liquid–vapor and ice–vapor curves for biphasic equilibria and the triple point at the intersection of these three curves (Fig. 7.28). The distinguishing feature is the negative slope of the curve for the ice–liquid equilibrium, which appears to be infinite. Considering an isothermal transformation that crosses this curve and is carried out in the direction of increasing pressures, from ice to liquid water, it is shown that ice melts when sufficient pressure is applied to it. Since increase of pressure always leads to increased compactness, one concludes that liquid water is denser than ice, that is, ice fluctuates on liquid water, and this is because the hexagonal channels of ice I_h produce a less dense structure than that of liquid water.

In the liquid–vapor equilibrium, the chemical potentials of the liquid and vapor are equal,

$$\mu_{\text{liq}}(T) = \mu_{\text{gas}}(p, T) \tag{7.27}$$

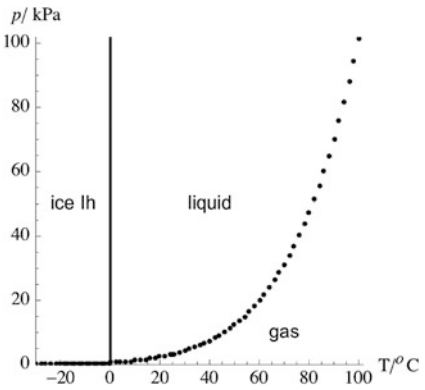


Fig. 7.28 Phase diagram of water for ordinary pressures (Haynes 2011). The triple point is at T = 273.16 K and p = 0.611657 kPa. The critical point (not shown) occurs at T = 647.10 K and p = 22064 kPa. This figure was obtained with *Mathematica*

where T is the boiling temperature at pressure p. From (7.27) and (1.149) (this expression gives the chemical potential of an ideal gas, an assumption that will soon be justified), we can write

$$\mu^0_{\text{liq}}(T) = \mu^0_{\text{gas}}(T) + RT \ln\left(\frac{p}{p^0}\right) \tag{7.28}$$

where p^0 is the reference pressure (p^0 = 1 bar = 100 kPa). Then,

$$\mu^0_{\text{gas}}(T) - \mu^0_{\text{liq}}(T) \approx RT\,\eta(T) \tag{7.29}$$

where

$$\eta(T) = \ln[p^0/p(T)] \tag{7.30}$$

At 298.15 K, p = 3.1699 kPa (Haynes 2011) and

$$\eta(298.15) = \ln\left(\frac{100}{3.1699}\right) \approx 3.5 \tag{7.31}$$

At 298.15 K, one has $RT \approx 2.5$ kJ mol^{-1}, and the difference between the standard chemical potentials of vapor and liquid water in equilibrium is approximately 9 kJ mol^{-1} ($\approx 2.5 \times 3.5$). Considering that the chemical potential is a partial molar Gibbs energy and the water vapor at such a low pressure (p = 3.1699 kPa) can be considered an ideal gas, (7.29) provides an estimate of the Gibbs energy necessary for vaporizing one mole of liquid water at 298.15 K.

The liquid–gas equilibrium curve has a terminal point called the **critical point** (not shown in Fig. 7.29) at T = 647.10 K and p = 22,064 kPa. Beyond the temperature of the critical point, the vapor cannot be liquefied by pressure increase. The critical point can be approached through a transformation that follows the liquid–

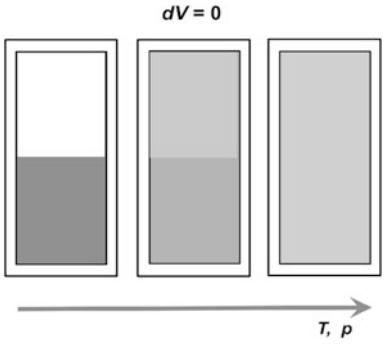


Fig. 7.29 Starting with the liquid–vapor equilibrium, an increase of temperature at constant volume leads the system to the critical point and beyond

vapor curve and is carried out at constant volume on a closed system consisting of liquid water in equilibrium with its vapor. As temperature increases, the pressure and the density of the vapor increase, while the density of the liquid decreases (the average density of liquid and vapor inside the sample is constant, since both the total mass of water and the volume are fixed). When approaching the critical point, the interfacial surface that distinguishes liquid from vapor gradually disappears, and an extremely dense vapor is formed called **supercritical fluid** (Fig. 7.29).

7.12 Water as Solvent

For a solution of one solute, it is possible to group the general types of intermolecular interactions as solute–solvent, solvent–solvent, and solute–solute. In a diluted solution of a nonionic solute, solute–solute interactions can in principle be neglected, since there is a small probability of two solute molecules being at an interaction range distance. However, for solutions of ionic compounds, the electrostatic interactions between positively and negatively charged ions may be important even for diluted solutions, since the potential energy of interaction between two point charges varies as the long-range function $1/R$, where R is the distance between the charges. For aqueous diluted solutions, the solvent concentration is approximately equal to 55.5 M ($\approx$1000 g dm^{-3}/18 g mol^{-1}). Considering a 0.10 M nonionic solution, for each solute molecule there are about 5.5×10^2 H_2O molecules. Solute–water interactions may appreciably affect the hydrogen bond network and the structure of water in a neighborhood of each solute molecule.

Solubility is the concentration of a saturated solution where the solute in solution is in equilibrium with the pure solute. The sodium chloride solubility in water at 25 °C is about 6.15 mol of sodium chloride per kilogram of water; that is, a saturated sodium chloride solution contains nine water molecules for each pair of sodium–chloride ions. At the same temperature, the solubility of carbon tetrachloride in water is about 0.0042 mol per kilogram of water, meaning that for each carbon tetrachloride molecule, there are more than 13,000 water molecules. Mixing carbon tetrachloride with water leads to distinct phases, and thus one usually says that carbon tetrachloride and water are immiscible liquids, that is, they do not (significantly) mix. On a molar basis, sodium chloride is approximately 1500 times more soluble in water than in carbon tetrachloride.

7.12.1 Electric Permittivity

We are now interested in finding a solvent property that may quantify the solvent polarity. Being a molecular substance, the solvent is a nonconducting material or dielectric. When we apply an electric field, charges are slightly displaced from their equilibrium positions, and, in the liquid phase, the total molecular dipoles (permanent plus induced dipoles) tend to become aligned with the applied electric field.

If the polarization is uniform, the displaced charges cancel in neighboring molecules and so appear only on the surfaces (Fig. 7.30). The induced electric field opposes the external electric field. Therefore, within the dielectric, the applied electric field is reduced, due to the internal electric field. The ability of the dielectric to store charge is called **capacitance**, denoted by C. For a parallel-plate **capacitor**, if the charges on the plates are $+q$ and $-q$ and the electric potential between the plates is V, the capacitance is given by

$$C = q/V \tag{7.32}$$

[the SI unit of capacitance is called the farad (F) after Michael Faraday (1791–1867); 1 F = 1 C V^{-1}]. The capacitance of a parallel-plate capacitor can be related to its geometry and the electric nature of the material between plates as

$$C = \varepsilon A/d \tag{7.33}$$

where A is the plate area, d is the plate separation, and ε is the **permittivity** of the material between the plates. The permittivity of vacuum is a universal constant (ε_o = 8.854187816 10^{-12} F m^{-1}). Usually, the permittivity of a dielectric is equated to the product of the permittivity of vacuum by a dimensionless factor called the **relative permittivity** ε_r, where

$$\varepsilon = \varepsilon_0 \varepsilon_r \tag{7.34}$$

The potential energy of interaction between two point charges is given by

$$E = \frac{q_1 q_2}{4\pi\varepsilon r} \tag{7.35}$$

where ε is the permittivity of the material in which the charges q_1 and q_2 are immersed, and r is the distance between them. Both in (7.33) and (7.35), the

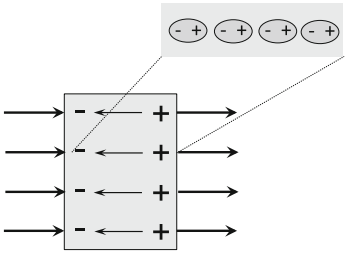


Fig. 7.30 An external electric field induces an electric field inside the dielectric. *Inset* shows details of induced dipoles in the molecules of the dielectric

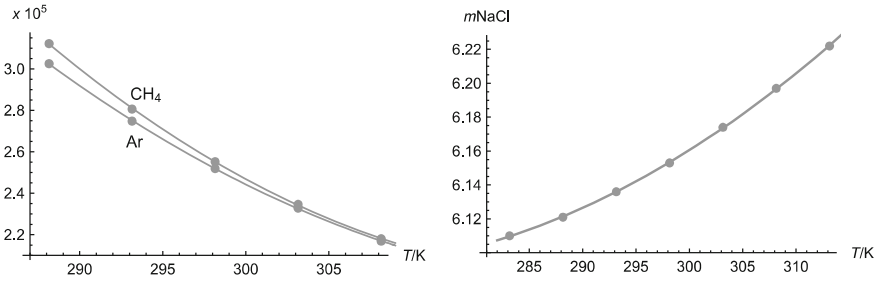


Fig. 7.31 Argon and methane solubilities in water given by mole fractions (note the 10^5 scaling factor) and the solubility of sodium chloride given as molality, as functions of temperature (Haynes 2011). This figure was obtained with *Mathematica*

permittivity can be seen as a scaling factor that measures the polarity of the medium.

At ambient temperature, the relative permittivity of water is 78.3553. This means that the potential energy of attraction between sodium and chloride ions in water is 78 times smaller than the potential energy of attraction for the same ions, at the same distance, in free space. Hence, water is quite efficient for promoting the dissociation of sodium chloride. In an aqueous sodium chloride solution, the local electric field created by the Na^+ and Cl^- ions reorients the neighboring water molecules around the ions by aligning their dipole moments with the electric fields generated by the ions. In contrast with water, the relative permittivity of carbon tetrachloride at ambient temperature is 2.2280. Since electrostatic interactions are much less important in a solvent of high permittivity than in a solvent of low permittivity, sodium chloride is very soluble in water but is insoluble in carbon tetrachloride.

7.13 Simple Nonpolar Solutes

Simple nonpolar solutes like argon and methane have extremely small solubilities in water that decrease with increasing temperature, in contrast with the large solubilities of electrolytes like sodium chloride that increase with temperature (Fig. 7.31). In order to understand these trends, we consider, on the one hand, the potassium ion–water molecule interactions, which are dominated by electrostatic interactions between the potassium ion charge and the dipole moment of neighboring water molecules (the electric field produced by the potassium ion aligns the dipole moments of surrounding water molecules, thus producing a solvated potassium ion, $[K(OH_2)_x]^+$), and on the other hand, the absence of charge and any electric moment in the argon atom, implying the nonexistence of electrostatic interactions between the argon atom and the water molecules, with these becoming available to form hydrogen bonds between them (Fig. 7.32). An increase in

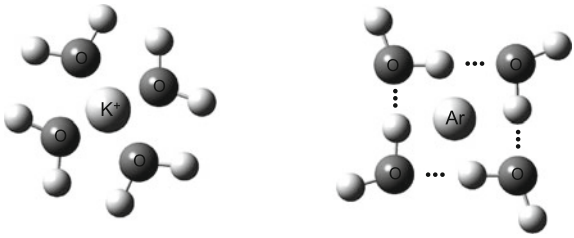


Fig. 7.32 Schematic representation of the solvation in water of a potassium ion and an argon atom

temperature leads to the disruption of the argon atom cavities by progressively breaking hydrogen bonds between water molecules on the surfaces of these cavities.

The argon and methane cavities are reminiscent of the convex polyhedra formed by hydrogen-bonded water molecules that include gas molecules and are the building blocks of solid structures known as gas hydrates. Being in the liquid phase, the mobility of the water molecules on the surfaces of the nonpolar solute cavities is expected to be large when compared with the reduced mobility of the hydrogen-bonded water molecules that form the convex polyhedra of gas hydrates.

The extremely low solubilities in water of simple nonpolar solutes like argon and methane make it possible to look at their aqueous solutions mainly as water subject to small perturbations on the hydrogen bond network of water (solvent–solvent interactions) around each solute molecule.

7.13.1 Ostwald Coefficient

Consider now a gas like argon or methane in contact with water. A decrease in the volume of the gas as a result of its dissolution in water, at a defined temperature and constant pressure of 1 atm (=101,325 Pa), corresponds to an irreversible process leading to equilibrium (Fig. 7.33). The ratio of the gas volume change ΔV_g to the volume of the aqueous solution V_{sln}, at equilibrium,

$$\gamma_s = \left(\frac{\Delta V_g}{V_{sln}}\right)_{eq} \tag{7.36}$$

is called the **Ostwald coefficient** and can be used to measure the gas solubility in water.

The Ostwald coefficient can also be defined as the ratio of the molar densities of the nonpolar solute in the liquid to gas phases at equilibrium,

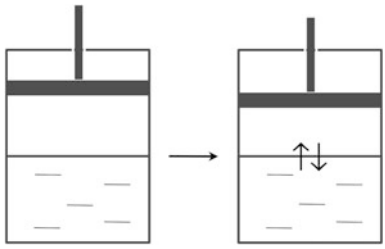


Fig. 7.33 Schematic representation of an irreversible process of dissolution in water of a nonpolar solute like argon or methane, at a constant pressure of 1 atm (=101,325 Pa) and defined temperature, leading to a decrease in the volume of the gas phase

$$\gamma_s = \left(\frac{\Delta V_g}{V_{sln}}\right)_{eq} = \left(\frac{n_s/V_{sln}}{n_s/\Delta V_g}\right)_{eq} = \left(\frac{\rho_{s,sln}}{\rho_g}\right)_{eq} \tag{7.37}$$

where n_s is the amount of the gas solute dissolved in water, and $\rho_{s,sln}$ is the molar density of the solute in the liquid phase. At equilibrium, the whole system has two phases, two independent components, and two degrees of freedom, which are temperature and pressure. The chemical potentials of the solute in the liquid and in the gas phases are equal,

$$\mu_{s,sln} = \mu_g \tag{7.38}$$

since the aqueous solution is saturated in the gas. The chemical potential of the gas, assumed to be ideal (the solute is nonionic and extremely diluted), is

$$\mu_g = \mu_g^0(T) + RT\ln\left(\frac{p}{p^0}\right) \tag{7.39}$$

and the gas pressure is

$$p = \rho_g RT \tag{7.40}$$

Substitution of (7.40) in (7.39) gives

$$\mu_g = \underbrace{\mu_g^0(T) + RT\ln\left(RT/p^0\right)}_{\mu_g^0(p^0,T)} + RT\ln\rho_g = \mu_g^0\left(p^0,T\right) + RT\ln\rho_g \tag{7.41}$$

We now consider the chemical potential of the solute in the liquid phase. The extremely low solubility of a nonpolar solute like argon or methane in water and, as a result, the random distribution of its molecules in water lead to an ideal solution. Hence, we can assume unitary activity coefficients and write

$$\mu_{s,sln} = \mu^0_{s,sln}(T) + RT \ln x_{s,sln} \tag{7.42}$$

where $x_{s,sln}$ is the solute mole fraction in the liquid phase (the solution). Considering the extremely low solubility of the nonpolar solute in water, then

$$x_{s,sln} \approx \frac{n_{s,sln}}{n_S} = \frac{n_{s,sln}/V_{sln}}{n_S/V_{sln}} = \frac{\rho_{s,sln}}{\rho_S} \tag{7.43}$$

where V_{sln} is the volume of the liquid phase, and n_S and ρ_S are the amount of solvent and the molar density of the solvent, respectively. Substitution of (7.43) in (7.42) leads to

$$\mu_{s,sln} = \underbrace{\mu^0_{s,sln}(T) - RT\ln\rho_S}_{\mu^0_{s,sln}(\rho_S,T)} + RT\ln\rho_{s,sln} = \mu^0_{s,sln}(\rho_S, T) + RT\ln\rho_{s,sln} \tag{7.44}$$

At equilibrium, the chemical potentials of the nonpolar solute in the liquid and gas phases are equal, that is,

$$\mu^0_{s,sln}(\rho_S, T) + RT\ln\left(\rho_{s,sln}\right)_{eq} = \mu^0_g\left(p^0, T\right) + RT\ln\left(\rho_g\right)_{eq} \tag{7.45}$$

Rearranging this equation and using the last equality of (7.37), we obtain

$$\mu^0_{s,sln}(\rho_S, T) - \mu^0_g\left(p^0, T\right) = -RT\ln\gamma_s \tag{7.46}$$

This equation shows that the Ostwald coefficient is a function of temperature. Considering the physical meaning of the chemical potential as a partial molar Gibbs energy, then the first member of (7.46) is the Gibbs energy for the transfer of one mole of the nonpolar gas from the gas to the liquid phase,

$$\Delta G_{g\to sln}(T) = -RT\ln\gamma_s(T) \tag{7.47}$$

where

$$\Delta G_{g\to sln}(T) \equiv \mu^0_{s,sln}(\rho_S, T) - \mu^0_g\left(p^0, T\right) \tag{7.48}$$

Using the Gibbs–Helmholtz equation [see (1.179)], we can write

$$\frac{\partial\left(\frac{1}{T}\Delta G_{g\to sln}\right)}{\partial T} = -\frac{\Delta H_{g\to sln}}{T^2} \tag{7.49}$$

Therefore, the temperature-dependence of the Ostwald coefficients allows one to determine the $\Delta G_{g\to sln}$ temperature-dependence and to find $\Delta H_{g\to sln}$ at each defined temperature. The corresponding entropy $\Delta S_{g\to sln}$ is then obtained from

$$\Delta G_{g\to sln} = \Delta H_{g\to sln} - T\Delta S_{g\to sln} \tag{7.50}$$

At ambient temperature (298.15 K), the approximate values of the thermodynamic state functions for argon and methane are as given in the following table.

Solute	$\Delta G_{g\to sln}$/ kJ mol^{-1}	$\Delta H_{g\to sln}$/ kJ mol^{-1}	$\Delta S_{g\to sln}$/ J mol^{-1} K^{-1}
Argon	8	–10	–60
Methane	8	–11	–65

Negative entropy values point to the formation of hydrogen bonds involving water molecules on the surfaces of the solute cavities. As heat is released due to the formation of these hydrogen bonds, enthalpy values become negative. Note that $\Delta H_{g\to sln}$ and $T\Delta S_{g\to sln}$ present approximately parallel temperature dependencies (Fig. 7.34). Since the negative values of $T\Delta S_{g\to sln}$ slightly exceed the negative values of $\Delta H_{g\to sln}$, the values of $\Delta G_{g\to sln}$ become positive and small.

7.13.2 Hydrophobic Interaction

We have been dealing with nonpolar solutes of small molecules that are gases when pure. However, when the pure nonpolar solute is a liquid and its solubility in water is exceeded, liquid phase separation occurs between water and the liquid nonpolar solute. Considering a gradual increase of the nonpolar solute, it is inevitable to conclude that the liquid nonpolar solute phase results from **coalescence of solute cavities** (Fig. 7.35). This is a gradual and progressive irreversible process that begins well before phase separation with the coalescence of solute cavities, with

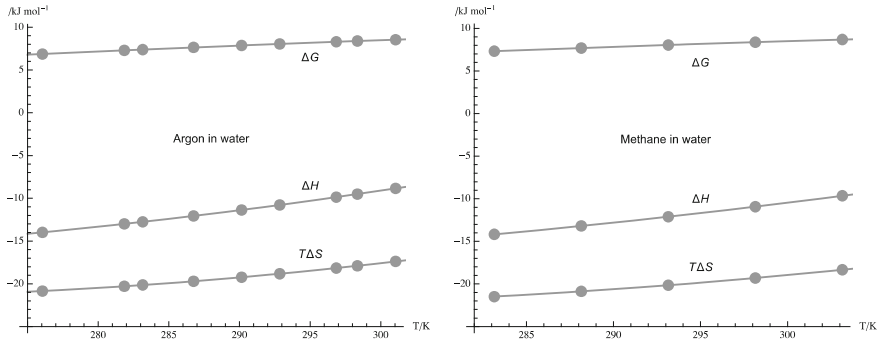


Fig. 7.34 Thermodynamic functions for the transfer of one mole of the nonpolar solute from the gas phase to the ideal solution, in the liquid phase, as functions of temperature [Ostwald coefficients for argon in water were taken from Ben-Naim and Baer (1963), and those for methane in water were taken from Yaacobi and Ben-Naim (1973)]. This figure was obtained with *Mathematica*

one solute molecule forming larger cavities with two solute molecules, followed by the formation of cavities with three solute molecules, four solute molecules, and so on, until large clusters of solute molecules can eventually form a distinct liquid phase.

A nonpolar solute in water that leads to liquid phase separation is called an **oil**. A typical example of an oil is *n*-decane, a linear alkane with ten carbon atoms per molecule. The process of coalescence of solute cavities cannot ignore the relevant role of water in the formation of hydrogen bonds between water molecules on the surfaces of the solute cavities (hydrogen bonds mainly oriented as tangents to the cavity surface), and in the important contribution that the high surface tension of water gives to the reduction of the surface area during coalescence of solute cavities. In fact, the ratio of surface area to volume of a cavity decreases as solute cavities coalesce (Fig. 7.35). In the absence of the guest molecule, solute cavities could not exist. They would become unstable and would collapse. The dispersion interaction between nonpolar solute molecules in the solute cavities is often called **hydrophobic interaction**.

7.14 Ionic Solutions

Due to its high relative permittivity water can easily dissociate strong electrolytes, yielding solutions that conduct electricity (ionic solutions). Consider the dissociation of a strong electrolyte

$$C_nA_m \rightarrow nC^{m+} + mA^{n-} \tag{7.51}$$

whose equation for electric neutrality is given by

$$n(m+) + m(n-) = 0 \tag{7.52}$$

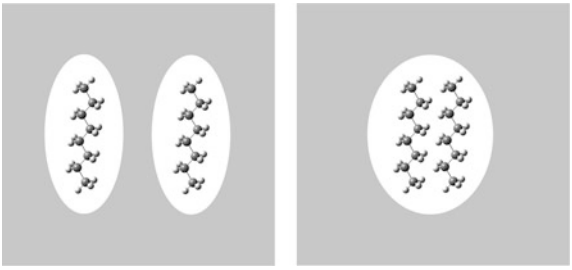


Fig. 7.35 Coalescence of oil-in-water cavities leads to a decrease in the surface area per oil molecule and to an increase in dispersion interactions between oil molecules

The individual activity coefficients of ions cannot be experimentally determined, since the solutions are electrically neutral, i.e., they contain positively charged ions and negatively charged ions that cannot be separated into different systems. However, in order to arrive at the chemical potential of the electrolyte (a measurable thermodynamic quantity), we begin with the chemical potentials of the cation and anion

$$\mu_{C^{m+}} = \mu^0_{C^{m+}} + RT\ln(\gamma_{C^{m+}} x_{C^{m+}}) \quad \mu_{A^{n-}} = \mu^0_{A^{n-}} + RT\ln(\gamma_{A^{n-}} x_{A^{n-}}) \tag{7.53}$$

For each mole of the electrolyte, the mole fractions of the cation and anion [see (7.51)] are given by

$$\text{cation:}\frac{n}{n+m} \quad \text{anion:}\frac{m}{n+m} \tag{7.54}$$

Applying these mole fractions to (7.53) and adding the resulting equations, we obtain the chemical potential for the electrolyte, given by

$$\mu_{salt} = \mu^0_{salt} + RT\ln x_{salt} + RT\ln\gamma_{salt} \tag{7.55}$$

where

$$\begin{aligned} \mu_{salt} &= \frac{n\mu_{C^{m+}} + m\mu_{A^{n-}}}{n+m} & \mu^0_{salt} &= \frac{n\mu^0_{C^{m+}} + m\mu^0_{A^{n-}}}{n+m} \\ x_{salt} &= \left(x^n_{C^{m+}} x^m_{A^{n-}}\right)^{1/(n+m)} & \gamma_{salt} &= \left(\gamma^n_{C^{m+}} \gamma^m_{A^{n-}}\right)^{1/(n+m)} \end{aligned} \tag{7.56}$$

Note that the arithmetic mean of the chemical potentials of the cation and anion corresponds to the geometric mean of mole fractions and activity coefficients, since these are arguments of a logarithm function. For sodium chloride, $m = 1$ and $n = 1$, and the mean activity coefficient is given by

$$\gamma_{salt} = \left(\gamma_{C^+}\gamma_{A^-}\right)^{1/2} \tag{7.57}$$

Most thermodynamic data express concentration variables in the molality scale. For this reason, we now convert mole fractions to molalities. Molality is the amount of solute per kilogram of solvent. Unlike molarities, molalities are not defined in terms of a volume of solution and so do not depend on temperature and are not affected by addition of other solutes to the solution. Starting with the definition of mole fraction of a solute, we can write, for a diluted solution,

$$x_s = \frac{n_s}{n_s + n_S} \approx \frac{n_s}{n_S} = \frac{n_s/1\,\text{kg solvent}}{n_S/1\,\text{kg solvent}} = \frac{m_s}{m_S} \tag{7.58}$$

where n_s and n_S are the amounts of solute and solvent and m_s and m_S are the solute and solvent molalities. For water as solvent of a diluted solution, $m_S \approx 55.55$ mol

(kg water)$^{-1}$. Note that $1/m_S \approx 0.018$ kg mol^{-1}, the molar mass of water expressed in kilograms (**E11**).

We have previously considered the chemical potential for a nonideal solution in terms of the activity of a solute

$$\mu(T) = \mu^0(T) + RT \ln a_s \tag{7.59}$$

where

$$a_s = \gamma_s x_s \tag{7.60}$$

Substitution of (7.58) in (7.60) and then in (7.59) and rearrangement of the resulting equation leads to

$$\mu(T) = \mu^0(T, m_S) + RT \ln(\gamma_s m_s) \tag{7.61}$$

where

$$\mu^0(T, m_S) = \mu^0(T) - RT \ln m_S \tag{7.62}$$

In an ionic solution, competition between thermal motion, which tends to randomly distribute ions in solution, and coulombic interaction, which attracts ions of opposite charge and repels ions with a charge of the same sign, leads to a slight imbalance that favors the coulombic attraction between ions of opposite charge. This is because the latter tend to stay around a particular ion longer than ions of the same charge, thus leading, over time integration, to a spherical charge distribution called **ionic atmosphere** (Fig. 7.36). Around a particular anion, the preponderance of cations over anions (anions are repelled) leads to a positively charged distribution that is the ionic atmosphere of the anion. In turn, around a particular cation, the preponderance of anions over cations (cations are repelled) leads to a negatively charged distribution that is the ionic atmosphere of the cation. Since the net charge within each ionic atmosphere is nearly zero, every macroscopic volume, no matter how small it is, is electrically neutral. The distance over which a charge q is shielded by the charge distribution of ions of opposite charge is called the **Debye length**.

The Debye–Hückel theory, formulated in 1923 by Peter Debye (1884–1966) and Eric Hückel (1896–1980), can evaluate departures from ideal behavior in ionic solutions. This theory works well for diluted solutions. The **Debye–Hückel limiting law** calculates the activity coefficient of an electrolyte at very low concentrations in water,

$$\log \gamma_{salt} = -A|z_c z_a| I^{1/2} \tag{7.63}$$

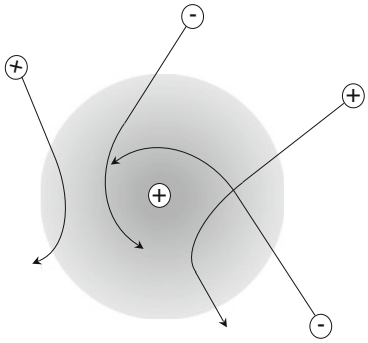


Fig. 7.36 Schematic representation of motion of ions around a particular cation. Note that anions are attracted by the cation, whereas cations are repelled. Therefore, anions remain longer than cations within the ionic atmosphere

where z_c and z_a are the charge numbers of the cation and anion, $A = 0.509$ for water (note that this value uses a base-10 logarithm), and I represents the **ionic strength** which is given by

$$I = \frac{1}{2}\sum_k z_k^2 \left(m_k/m^0\right) \tag{7.64}$$

where the summation is over all ions in solution, z_k is the charge number of ion k, m_k is its molality, and $m^0 = 1$ mol (kg water)$^{-1}$ (**E12**). The following figure presents the mean activity coefficients of sodium chloride, potassium chloride, and hydrogen chloride as functions of the solute molality, and shows that only extremely diluted ionic solutions have activity coefficients equal to 1 (Fig. 7.37).

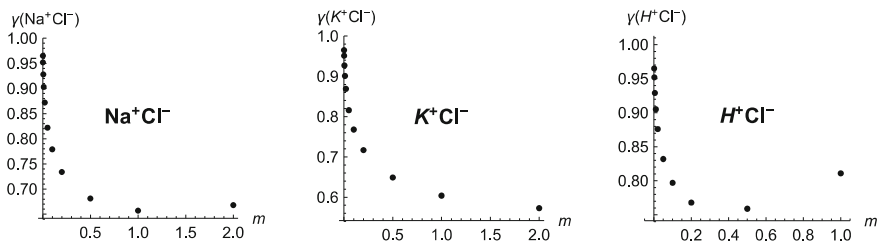


Fig. 7.37 Mean activity coefficients of sodium chloride, potassium chloride, and hydrogen chloride as functions of solute molality in aqueous solutions. Data taken from Haynes (2011). This figure was obtained with *Mathematica*

7.15 Amphipathic Molecules

Sodium decanoate, $CH_3(CH_2)_8COO^-\ Na^+$, has a carboxylate group, $–COO^-$, bonded to a linear alkyl group with nine carbon atoms $CH_3(CH_2)_8$. These molecular fragments have distinct polarity features that can be illustrated by the electrostatic potential of the decanoate ion mapped over an isodensity surface (isovalue = 0.00040; Fig. 7.38). The carboxylate group, a hydrophilic fragment, has a strongly negative electrostatic potential due to its negative charge equally distributed over the oxygen atoms. In turn, the alkyl chain, a hydrophobic molecular fragment, has an electrostatic potential mostly represented by the green color that corresponds to low or null electrostatic potentials, a feature that is clearly observed after the third carbon atom, since the first three carbon atoms are under the close influence of the carboxylate ion. Molecules that exhibit both hydrophilic and hydrophobic properties are called **amphipathic molecules** (*amphi* = both + Greek *pathikos*, from *pathos* = experience).

How do amphipathic molecules organize in water? When a solid formed by amphipathic molecules is gradually added to water in small amounts, the surface tension is one of the properties that directly reflects the course of molecular events in the medium, with the amphipathic molecule having a clear preference for being adsorbed on the liquid–air and liquid–solid interfaces. In this way, the amphipathic molecule prevents its hydrophobic fragment from being exposed to water, to which it directs its hydrophilic group (Fig. 7.39). Therefore, as the concentration of the amphipathic molecule gradually increases, the surface tension of the mixture decreases with respect to the surface tension of pure water (71.99 mJ m^{-2} at 25°C), and for this reason, the amphipathic solute is generally called a **surfactant** (see Vargaftik et al. 1983). In the strict technical sense, a surfactant is a compound that for a concentration not greater than 0.01 M, lowers the water surface tension by 30 mJ m^{-2} or more (see Berg 2010). As soon as the liquid–air and liquid–solid interfaces become totally covered by the surfactant the progressive increase in its concentration ceases to cause appreciable change on the surface tension of the medium, and the graph of the surface tension versus the initial concentration of the surfactant exhibits an approximately horizontal pattern (Fig. 7.39). Since the

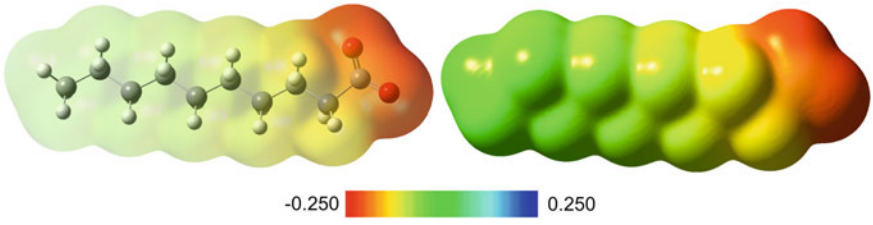


Fig. 7.38 Electrostatic potential mapped on isodensity surface with isovalue 0.00040 for the decanoate ion [*Gaussian* 09 B3LYP/6-31G(d,p) calculation]. On the *left*, the isodensity surface was intentionally made transparent to visualize the molecular geometry

surfactant adsorption to the interfaces has reached a maximum, the progressive increase in the surfactant leads to organized structures inside the medium, as the amphipathic molecules find a new way of preventing their hydrophobic fragments from being exposed to water. This time, they form organized structures where the hydrophilic groups are exposed to water and the hydrophobic chains take shelter inside organized structures called **micelles** (Fig. 7.39). The surfactant initial concentration at which micelles are formed is called **critical micelle concentration**, CMC. The latter corresponds to an energy minimum that results from a balance between attractive dispersion interactions between the tails of the amphipathic molecules inside the micelle and the repulsive electrostatic interaction involving the heads of the amphipathic molecules. A deep energy minimum leads to a defined critical micelle concentration and a specific aggregation number. Experimental kinetic studies reveal two kinds of relaxation processes involving micelles: one involves fast exchange of monomers between micelles and the aqueous phase and corresponds to a relaxation time of order 10^{-6} s; the other is the dissociation of micelles into monomers, which has a relaxation time of order 10^{-3} s (Oh and Shah 1993).

Let us assume the following equilibrium:

$$n\mathrm{M}_1 \rightleftarrows \mathrm{M}_n \tag{7.65}$$

where M_1 represents a monomer and M_n stands for a micelle with aggregation number equal to n. If C_0 is the total concentration of monomers, C_1 is the concentration of monomers in aqueous phase, C_n is the concentration of monomers in the micelles and no equilibrium other than (7.65) exists, then

$$C_0 = C_1 + C_n \tag{7.66}$$

Dividing both members by C_0 leads to

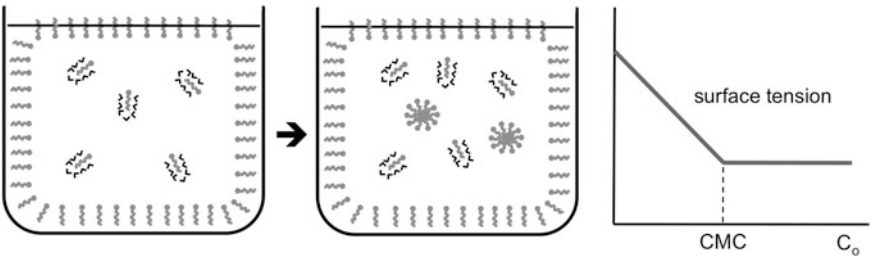


Fig. 7.39 As an amphipathic solid is gradually added to water, the amphipathic molecules begin by covering the interfacial liquid–air and liquid–solid surfaces of the beaker, and the surface tension of the medium decreases. Once the critical micellar concentration is reached, micelles are formed, and the surface tension shows an approximately horizontal variation (note that the graph is merely schematic)

$$1 = x_1 + x_n \tag{7.67}$$

where x_1 is the mole fraction of the monomers in the bulk solvent (=C_1/C_0) and x_n is the mole fraction of the monomer in the micelles (=C_n/C_0). Under ideal conditions (negligible monomer–micelle interactions), the equilibrium constant for (7.65) is given by

$$K_n = \frac{(C_n/n)}{C_1^n} \tag{7.68}$$

where the numerator is the concentration of micelles. Substitution of

$$C_n = x_n C_0 \quad C_1 = (1 - x_n)C_0 \tag{7.69}$$

in (7.68) gives

$$K_n = \frac{x_n}{n(1 - x_n)^n} C_0^{1-n} \tag{7.70}$$

After taking the $1/n$ root of both members and rearranging, we obtain

$$x_n = 1 - C_0^{(1/n-1)} \frac{x_n^{1/n}}{(nK_n)^{1/n}} = f\left[C_0, x_n^{1/n}\right] \tag{7.71}$$

This equation cannot be algebraically solved with respect to x_n, since this variable appears in both members. Therefore, it needs to be solved numerically by iteration, where an initial guess of x_n is introduced in the second member of (7.71), and this member of the equation is solved to give an improved value of x_n in the first member of (7.71). This value is again introduced in the second member, leading to a new and presumably better value of x_n. Then, we can write

$$x_{n(i+1)} = f\left[C_0, x_{n(i)}^{1/n}\right] \tag{7.72}$$

For n = 20 and $K_n = 10^{2n}$, the graph of x_n as a function of C_0/mM is presented in Fig. 7.40 (see *Mathematica* code **M6** for details of the iteration process) and shows that the mole fraction of monomers in the micelles, x_n, exhibits an initial abrupt increase for small values of C_0, and for C_0 greater than 200 mM, x_n enters a platform with small variations. The *Mathematica* code **M7** shows how to obtain an estimate of the critical micelle concentration (CMC) by determining C_0 through the interception of tangents to the curve (Fig. 7.40).

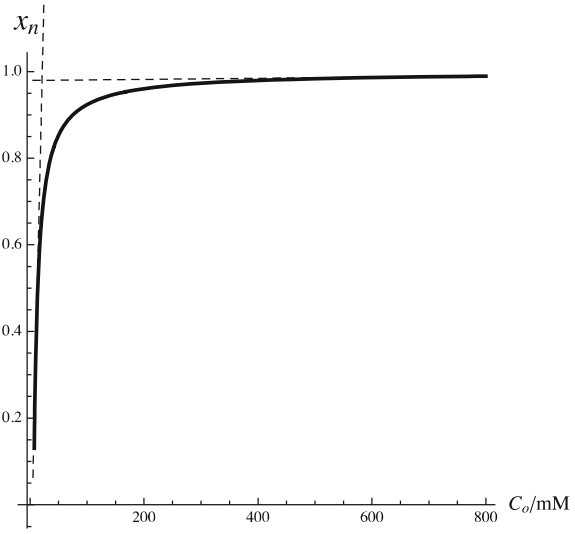


Fig. 7.40 Single equilibrium between n monomers and one micelle for $n = 20$ and $K_n = 10^{40}$, with the CMC value approximately equal to 21 mM. This figure was obtained with *Mathematica* (see *Mathematica* code **M3**)

7.15.1 Sodium Decanoate Micelles

For $C_0 >$ CMC, sodium decanoate in water forms spherical micelles with a more or less defined aggregation number. The methyl protons of the decanoate ion in the bulk solvent and in the micelles "feel" distinct environments that can be probed using ^{1}H nuclear magnetic resonance. Nuclear magnetic resonance (NMR) spectroscopy measures frequencies and intensities of spin transitions of nuclei with nonnull spin quantum number. In order to observe a nuclear spin transition, a magnetic field has to be applied to split the otherwise degenerate spin states. The chemical shift of a nucleus is given by

$$\delta = \frac{(\nu - \nu_{\text{ref}})}{\nu_{\text{ref}}} \times 10^6 \tag{7.73}$$

where ν is the frequency of the considered nucleus and ν_{ref} is the frequency of a reference, that for ^{1}H NMR is usually tetramethylsilane, $Si(CH_3)_4$. Since the order of magnitude of the frequency difference $\nu - \nu_{\text{ref}}$ is 10^{-6} ν_{ref}, one usually includes in (7.73) a factor 10^6, thus expressing the chemical shift in parts per million (ppm). The chemical shift of a nucleus is independent of the applied magnetic field, since both $\nu - \nu_{\text{ref}}$ and ν_{ref} are proportional to the applied magnetic field.

Consider now the single equilibrium (7.65) between the monomers of the decanoate ion in the bulk (D_2O) and the decanoate ion micelles. Our model system (see Andrade-Dias et al. 2007) consists of two proton states, one for the methyl group protons of the decanoate ion in the bulk, the other for the methyl group protons of the decanoate ion in the micelles. Nuclear magnetic resonance cannot distinguish the chemical shifts of the methyl group protons in both of these

environments, due to fast exchange in the NMR time scale between the monomer in the bulk and in the micelles. Hence, the observed chemical shift is the average of the chemical shifts of the methyl group protons in both of these environments, weighted by the corresponding mole fractions (fast exchange regime),

$$\delta = x_1\delta_1^0 + x_n\delta_n^0 \tag{7.74}$$

where δ_1^0 is the chemical shift for the methyl protons when $x_1 = 1$, i.e., when all the monomers are in the bulk solvent, and δ_n^0 is the chemical shift for the methyl protons when $x_n = 1$, i.e., when all the monomers are in the micelles Substitution of (7.67) in (7.74) and rearrangement of the obtained equation leads to

$$\Delta\delta = x_n\Delta\delta^0 \tag{7.75}$$

where

$$\Delta\delta = \delta - \delta_1^0 \quad \Delta\delta^0 = \delta_n^0 - \delta_1^0 \tag{7.76}$$

Substituting the first equality of (7.69) in (7.75) leads to

$$C_0\Delta\delta = C_n\Delta\delta^0 \tag{7.77}$$

For $C_0 <$ CMC, micelles do not occur and $C_n = 0$, that is, the first member of (7.77) is equal to zero. For $C_0 >$ CMC, the concentration of the monomers in the micelles (C_n) increases linearly with C_0, since $\Delta\delta^0$ is a constant in the experiment.

Figure 7.41 shows $C_0\,\Delta\delta$ versus C_0 for sodium decanoate in D_2O (D = ^{2}H). It can be seen that the general pattern predicted by Eq. (7.77) is approximately followed. The addition of an ionic compound like sodium fluoride increases the ions

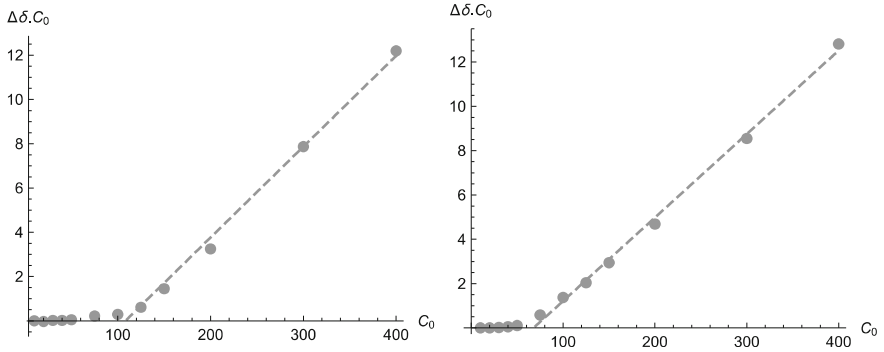


Fig. 7.41 $\Delta\delta.C_0$ versus C_0/mM for sodium decanoate in deuterated water (*left*), and with sodium fluoride ([NaF] = 300 mM) added (*right*). The intercepts of straight lines with the C_0 axis give approximate values for the critical micelle concentrations (*Left*, CMC≈110 mM; *Right*, CMC ≈ 73 mM). This figure was obtained with *Mathematica*

of opposite charge in the medium, thus reducing the repulsion between the polar heads of the amphipathic molecules in the micelles and lowering the critical micelle concentration.

7.16 Acids and Bases

When ammonia is in contact with hydrogen chloride, a white cloud of ammonium chloride is obtained that can be deposited as a white solid over a smooth cold surface. Ammonia is a gas, whose normal boiling point is $T_b = -33$ °C; hydrogen chloride is another gas, whose normal boiling point is $T_b = -85$ °C, and ammonium chloride is an ionic solid formed by NH_4^+ and Cl^- ions, with lattice energy ≈ -712 kJ mol^{-1} (see Kapustinskii 1956). The chemical reaction

$$HCl(g) + NH_3(g) \rightarrow NH_4^+ Cl^-(s) \tag{7.78}$$

corresponds to the transfer of a hydrogen ion from one HCl molecule that converts itself into a Cl^- ion to one NH_3 molecule that produces an ammonium ion NH_4^+. This reaction illustrates the concepts of acid and base that were developed by J.N. Bronsted (1879–1947, Danish) and T.M. Lowry (1874–1936, British) in 1923: the **acid** is the chemical species that donates one hydrogen ion; the **base** is the chemical species that accepts one hydrogen ion.

Let us substitute ammonia by water. In contrast to (7.78), the reactants now have different physical states, since pure hydrogen chloride is a gas and water is a liquid. The greater density of water "tilts" all molecular events on its side. In a 0.10 M aqueous solution, for each mole of solute there are about 5.5×10^2 H_2O molecules! Water has the important role of being the solvent, intervening in all chemical processes that occur in solution. The chemical equation is

$$HCl(g) + H_2O(aq) \rightarrow H_3O^+(aq) + Cl^-(aq) \tag{7.79}$$

where HCl acts as a Bronsted–Lowry acid by donating a hydrogen ion to one H_2O molecule and leaving behind a chloride ion, and H_2O acts as a Bronsted–Lowry base by receiving one hydrogen ion to form one hydronium ion, H_3O^+ (this ion has a triangular pyramidal geometry). The symbol aq in curved brackets indicates that the chemical species to which it applies is solvated by water. Since each hydrogen chloride molecule is surrounded by water molecules, each hydrogen chloride molecule does not need to migrate to form an acid–base pair.

For aqueous solutions with concentration not exceeding 1 M, scientific studies have concluded that an acid–base reaction includes H_2O molecules for solvation of the acid and/or the base, that is, H_2O molecules mediate the hydrogen ion transfer between acid and base. When mediated by one H_2O molecule, the hydrogen ion transfer becomes one order of magnitude slower ($\approx$1 ps) than the unmediated transfer. In turn, concentrated aqueous solutions present a bimodal mechanism,

since acid–base pairs also occur in solution, leading to the direct transfer of a hydrogen ion between acid and base (see Rini et al. 2003; Mohammed et al. 2005).

7.16.1 Autoionization of Water

The water molecule has both acid and base functionalities. The following chemical equation describes the **autoionization** equilibrium of water:

$$H_2O(aq) + H_2O(aq) \rightleftarrows H_3O^+(aq) + OH^-(aq) \tag{7.80}$$

From the thermodynamic point of view, every chemical species requires the definition of a standard state. However, there are no substances formed by ions with the same charge, since the extremely high electrostatic repulsion between the ions would prevent the formation of a stable substance. In aqueous solution, it is established by definition that at all temperatures,

$$\Delta H_f^0[H^+(aq)] \equiv 0 \quad S^0[H^+(aq)] \equiv 0 \quad \Delta G_f^0[H^+(aq)] \equiv 0 \tag{7.81}$$

and as a consequence,

$$C_p[H^+(aq)] = 0 \tag{7.82}$$

where C_p is the isobaric heat capacity. Since (7.81) and (7.82) are thermodynamically defined for the hydrogen ion H^+, thermodynamics represents the autoionization of water by the chemical equation

$$H_2O(aq) \rightleftarrows H^+(aq) + OH^-(aq) \tag{7.83}$$

The following table presents the thermodynamic quantities of the chemical species present in (7.83), at 298.15 K (see Haynes 2011).

T = 298.15 K	ΔH_f^0/kJ mol^{-1}	ΔG_f^0/kJ mol^{-1}	S^0/J mol^{-1} K^{-1}
H^+(aq)	0	0	0
OH^-(aq)	−230.0	−157.2	−10.9
H_2O(L)	−285.8	−237.1	70.0

Considering (7.81) and (7.83), we can write

$$\Delta G_w^0(T) = \left[\mu_{H^+}^0(T) + \mu_{OH^-}^0(T)\right] - \mu_{H_2O}^0(T) = -RT \ln K_w(T) \tag{7.84}$$

where K_w represents the **ionization constant** of water,

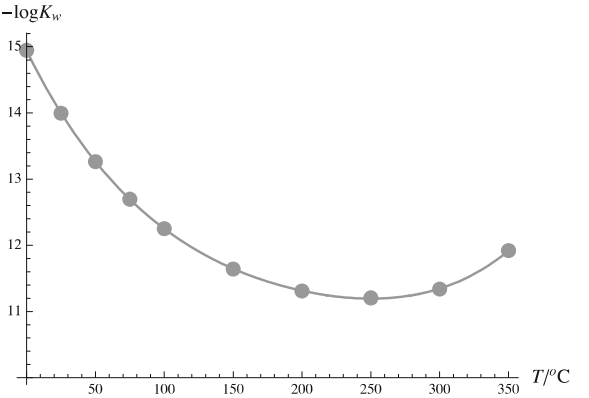


Fig. 7.42 pK_w as a function of temperature, in the range 0–350 °C [for the data used, see Bandura and Lvov (2006)]. This figure was obtained with *Mathematica*

$$K_{\mathrm{w}} = \frac{a_{\mathrm{H}^+} a_{\mathrm{OH}^-}}{a_{\mathrm{H_2O}}} \approx a_{\mathrm{H}^+} a_{\mathrm{OH}^-} \tag{7.85}$$

where a_k is the activity of chemical species k (note that $a_{\mathrm{H_2O}} \approx 1$). At 25 °C, the constant of autoionization of water is approximately equal to 1.0×10^{-14}. Figure 7.42 shows the temperature dependence of -logK_w, from 0 to 350 °C. The minimum of this curve occurs at 250 °C, with $-\log K_{\mathrm{w}} = 11.2$.

Since H_3O^+(aq) is equivalent to H^+(aq) + H_2O(aq) from the stoichiometric point of view, the activity of the hydronium ion is defined by

$$a_{\mathrm{H_3O^+}} = a_{\mathrm{H}^+} a_{\mathrm{H_2O}} \tag{7.86}$$

This definition implies that the constant of the equilibrium

$$\mathrm{H_3O^+} \rightleftarrows \mathrm{H_2O} + \mathrm{H^+} \tag{7.87}$$

is 1, which, in turn, means that H_3O^+(aq) and H^+(aq) + H_2O(aq) are also equivalent from the thermodynamic point of view. Note that the hydronium species H_3O^+ exists in water as such and as chemical species resulting from solvation with different numbers of water molecules, such as $H_5O_2^+$, $H_7O_3^+$, and $H_9O_4^+$. On the other hand, H^+ is a proton that really exists only in the transition state of an acid–base reaction like (7.78). From (7.86) and $a_{\mathrm{H_2O}} \approx 1$, it can be concluded that the equilibrium constant for the autoionization of water as expressed by (7.80) is numerically equal to the equilibrium constant for the autoionization of water obtained from (7.83).

7.16.2 Acid Ionization Constant

Consider now a weak acid HA and the ionization equilibrium in water

$$HA(aq) + H_2O(aq) \rightleftarrows H_3O^+(aq) + A^-(aq) \tag{7.88}$$

This chemical equation is equivalent to

$$HA(aq) \rightleftarrows H^+(aq) + A^-(aq) \tag{7.89}$$

whose thermodynamic equilibrium constant is given by

$$K_{HA} = \frac{a_{A^-} a_{H^+}}{a_{HA}} \tag{7.90}$$

Applying logarithms to both members and multiplying by −1, we obtain

$$-\log K_{HA} = -\log a_{H^+} + \log \frac{a_{HA}}{a_{A^-}} \tag{7.91}$$

The hydrogen ion activity is usually expressed by the solution's pH, defined by

$$pH \equiv -\log a_{H^+} \tag{7.92}$$

In addition, the acid pK_{HA} is defined by

$$pK_{HA} \equiv -\log K_{HA} \tag{7.93}$$

For very diluted aqueous solutions, we can assume that the activity coefficients are approximately equal to 1, and the substitution of (7.92) and (7.93) in (7.91) leads to

$$pH \approx pK_{HA} - \log \frac{x_{HA}}{x_{A^-}} \tag{7.94}$$

where the mole fractions satisfy the following equation:

$$x_{HA} + x_{A^-} = 1 \tag{7.95}$$

Substitution of (7.95) in (7.94) leads to

$$pH \approx pK_{HA} - \log \frac{(1 - x_{A^-})}{x_{A^-}} \tag{7.96}$$

Rearranging (7.96) and using (7.95), we obtain

$$x_{A^-} \approx \frac{1}{1 + 10^{(pK_{HA} - pH)}} \quad x_{HA} \approx \frac{1}{1 + 10^{(pH - pK_{HA})}} \tag{7.97}$$

Figure 7.43 shows, at left, x_{HA} and x_{A^-} as functions of pH for $pK_{HA} = 4$, and x_{HA} as a function of pH for $pK_{HA} = 2$, 3, 4, 5, and 6; x_{HA} and x_{A^-} are S-shaped functions of pH that sum to 1 and cross at the function value 1/2 when $pH = pK_{HA}$.

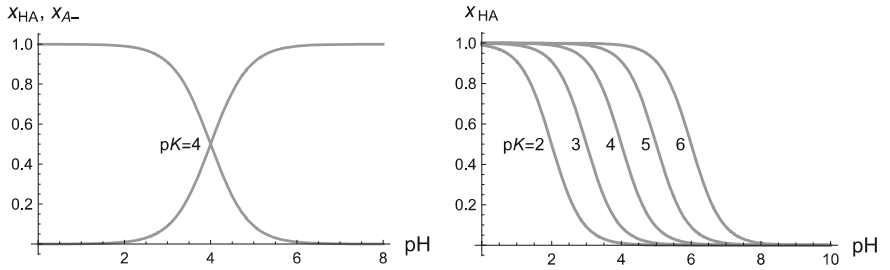


Fig. 7.43 x_{HA} and x_{A^-} as functions of pH for $pK_{HA} = 4$, at *left*, and x_{HA} as a function of pH, for $pK_{HA} = 2$, 3, 4, 5, and 6, at *right*. This figure was obtained with *Mathematica*

The way a solution resists a change in pH can be evaluated by dx_{A^-}/dpH. Figure 7.44 shows, at left, the pH as a function of x_{A^-} for $pK_{HA} = 4$ (this illustrates a **universal titration curve** since x_{A^-} is not affected by volume change, as occurs when the independent variable in a regular titration is the added volume of the base) and dx_{A^-}/dpH as a function of x_{A^-}, at right [note that dx_{A^-}/dpH is the inverse of the derivative of pH = $f(x_{A^-})$; see *Mathematica* code **M8**]. It can be concluded that the solution attains maximum buffer capacity at $x_{A^-} = 0.5$, when HA and A^- have the same mole fractions.

We now consider (7.90) and convert mole fractions in molalities [see (7.58)], obtaining

$$K_{HA} = \frac{a_{A^-} a_{H^+}}{a_{HA}} = \frac{x_{A^-} x_{H^+}}{x_{HA}} \frac{\gamma_{A^-} \gamma_{H^+}}{\gamma_{HA}} = \frac{m_{A^-} m_{H^+}}{m_{HA} m_S} \frac{\gamma_{A^-} \gamma_{H^+}}{\gamma_{HA}} = \frac{K_a}{m_S} \frac{\gamma_{A^-} \gamma_{H^+}}{\gamma_{HA}} \tag{7.98}$$

where

$$K_a = \frac{m_{A^-} m_{H^+}}{m_{HA}} \tag{7.99}$$

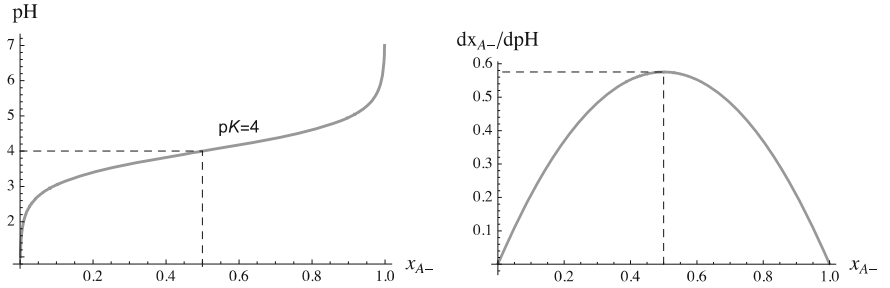


Fig. 7.44 pH as a function of x_{A^-} for $pK_{HA} = 4$, at *left*, and dx_{A^-}/dpH as a function of x_{A^-} with maximum at $x_{A^-} = 0.5$, at *right*. This figure was obtained with *Mathematica*

is called the **acidity constant** or **acid ionization constant**. Chemical species involved in thermodynamic expressions and in table values usually have their concentration expressed in molality. The thermodynamic values for the formation of ions in water at 298.15 K are referred to a standard state whose concentration is 1 molal.

7.16.3 Lewis Acids and Bases

We now consider chemical reactions that retain the essential characteristics of an acid-base reaction, yet involve acids that may not contain ionizable hydrogen atoms. The chemical reaction between ammonia and trimethylborane is one of such reactions and results in the formation of a compound of formula $(CH_3)_3B–NH_3$,

$$(CH_3)_3B + NH_3 \rightarrow (CH_3)_3B{-}NH_3 \tag{7.100}$$

Trimethylborane is a gas consisting of $B(CH_3)_3$ molecules, with normal boiling point $T_b = -20.2$ °C. The boron atom is electron-deficient, i.e., it is an electrophile. By contrast, the nitrogen atom of ammonia is a nucleophile. The complementary features of boron and nitrogen atoms cancel out when the compound $(CH_3)_3B\text{-}NH_3$ is formed. In the trimethylborane molecule, the boron atom lies in the plane defined by the carbon atoms, whereas in the compound $(CH_3)_3B\text{-}NH_3$, the boron atom and the carbon atoms form a triangular pyramid.

The elementary process of addition of a hydrogen ion to the nitrogen atom of ammonia,

$$H^+ + NH_3 \rightarrow {NH_4}^+ \tag{7.101}$$

shares with (7.100) the essential features of an acid–base chemical reaction: both the boron atom of trimethylborane and the hydrogen ion are electrophiles (the hydrogen ion is an extremely strong electrophile) that bond to ammonia through the lone electron pair of nitrogen. Hence, (7.100) is an acid–base reaction and trimethylborane an acid, despite lacking any ionizable hydrogen atom.

Gilbert Lewis (1875–1946), known for his association with the valence electron pair structures of molecules, captured the similarity between reactions (7.100) and (7.101) and proposed an acid–base concept that comprises both of these reactions. According to Lewis, an acid is a chemical species that contains an electrophilic center and attracts a nonbonding electron pair from the base (the nucleophile), thus forming a covalent bond between the electrophilic and the nucleophilic atoms or ions. Hence, (7.100) is a Lewis acid –base reaction, where $B(CH_3)_3$ is the **Lewis acid**, NH_3 is the **Lewis base**, and $(CH_3)_3B\text{-}NH_3$ is the **Lewis complex**.

The elementary process of addition of one hydrogen ion (a Lewis acid) to the oxygen atom of a water molecule (a Lewis base),

$$H^+ + H_2O \rightarrow H_3O^+ \tag{7.102}$$

leads to formation of one hydronium ion (a Lewis complex). This elementary acid–base reaction is similar to (7.101).

A different example of a Lewis acid–base reaction is given by a carbon dioxide molecule reacting with the hydroxide ion to form the hydrogenocarbonate ion,

$$CO_2(aq) + OH^-(aq) \rightarrow HCO_3^-(aq) \tag{7.103}$$

In this example, carbon dioxide is the Lewis acid, the hydroxide ion is the Lewis base, and the hydrogenocarbonate ion is the Lewis complex.

The solvation of a metal cation can also illustrate a Lewis acid–base reaction in which the Lewis acid is the metal cation, the coordinating water molecules are the Lewis bases, and the obtained aquo-complex is the Lewis complex. For instance,

$$Cu^{2+} + 6H_2O \rightarrow \left[Cu(H_2O)_6\right]^{2+} \tag{7.104}$$

Lewis and Bronsted–Lowry acid–base reactions are of different types: while the first is an **addition reaction** (the Lewis acid reacts with the Lewis base to form the Lewis complex), the second is a **displacement reaction** (the hydrogen ion is transferred from the Bronsted–Lowry acid to the Bronsted-Lowry base). According to Lewis nomenclature, the Bronsted–Lowry acid is the Lewis complex of a hydrogen ion (a Lewis acid) with a Lewis base, and a Bronsted–Lowry acid–base reaction is a chemical reaction of the following general kind:

$$\text{complex}_L + \text{base}_L \rightarrow \text{base}'_L + \text{complex}'_L \tag{7.105}$$

where the subscript L stands for "Lewis."

The Lewis acid concept is more general than the Bronsted–Lowry acid concept. Considering the base, there are no differences between the Lewis and the Bronsted–Lowry concepts: both the Lewis and the Bronsted–Lowry bases are nucleophiles that have at least one pair of nonbonding electrons in one of their electronegative atoms (typically, N, O, F). However, while the Bronsted–Lowry acid and the Lewis acid are both electrophiles, the Lewis acid may not have an ionizable hydrogen atom.

7.17 Standard Electrode Potentials

Thestandardelectrode potential defined to be zero is the **standard hydrogen electrode potential**, which corresponds to the half-reaction

$$H^+(aq) + e^- \rightarrow \frac{1}{2}H_2(g) \tag{7.106}$$

where the molality of the hydrogen ion is given by $m(H^+) = 1$ mol (kg of solvent)$^{-1}$, and the pressure of molecular hydrogen by $p(H_2) = 101.325$ kPa (Fig. 7.45). The thermic coefficient for the standard hydrogen electrode potential, $d\varepsilon^o/dT$, is also zero by definition, meaning that the above definition of the standard hydrogen electrode potential applies to all temperatures.

The International Union of Pure and Applied Chemistry (IUPAC) proposed a method for determination of hydrogen ion activity that uses the Harned cell,

$$\frac{1}{2}H_2(g) + AgCl(s) \rightarrow Ag(s) + H^+(aq) + Cl^-(aq) \tag{7.107}$$

(see Covington 2011; Haynes 2011). This cell involves the silver chloride–silver electrode

$$AgCl(s) + e^- \rightarrow Ag(s) + Cl^-(aq) \tag{7.108}$$

and the hydrogen electrode (7.106) and does not use any liquid junction. The method measures the potential difference between two Harned cells designed in such way that the obtained value does not depend on the standard hydrogen electrode potential, which is zero by definition, since this electrode potential is canceled in the potential difference. The measured potential difference depends only on the hydrogen ion activity of the hydrogen chloride aqueous solution in one of the cells in which HCl is present at a fixed molality [e.g., $m = 0.01$ mol (kg of solvent)$^{-1}$], and the mean activity coefficient is known (e.g., at 298.15 K, $\gamma_\pm = 0.904$). The method enables one to determine the hydrogen ion activity by the solution pH [see (7.92)] at different temperatures.

Half-equations (7.106) and

$$H_2O(aq) + e^- \rightarrow \frac{1}{2}H_2(g) + OH^-(aq) \tag{7.109}$$

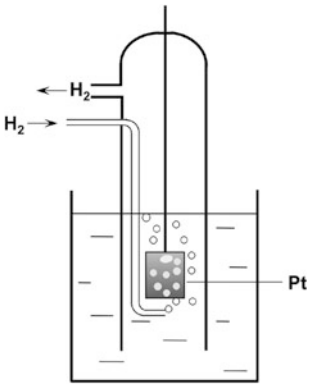


Fig. 7.45 Gaseous hydrogen adsorbed on the surface of the platinum plate of a hydrogen electrode. Platinum acts as a catalyst

involve the same oxidation states of hydrogen (+1, 0), in acid and alkaline solutions, respectively. Since the activities of H^+ and OH^- are interdependent by the ionization constant of water, the above half-equations are equivalent from the physical point of view, that is, they correspond to the same electrode potential, whether it is obtained by half-equation (7.106),

$$\varepsilon = -\frac{RT}{F}\ln\frac{a_{H_2}^{1/2}}{a_{H^+}} \tag{7.110}$$

or by half-equation (7.109),

$$\varepsilon = \varepsilon^0_{H_2O/H_2} - \frac{RT}{F}\ln\frac{a_{OH^-}a_{H_2}^{1/2}}{a_{H_2O}} \tag{7.111}$$

The equality between the second members of these equations leads to

$$\varepsilon^0_{H_2O/H_2} = \frac{RT}{F}\ln K_w \tag{7.112}$$

where K_w is the ionization constant of water. At 298.15 K, the standard hydrogen electrode potential in alkaline solution is $\varepsilon^0_{H_2O/H_2} = -0.828\,V$ (**E13**).

The following table presents several standard electrode potentials at 298.15 K (25 °C) and a pressure of 101.325 kPa (1 atm).

Standard electrode potentials (Values taken from Haynes 2011)

	Cl_2/Cl^-	Ag^+/Ag	Fe^{3+}/Fe^{2+}	Cu^+/Cu	Cu^{2+}/Cu	$AgCl/Ag+Cl^-$	H^+/H_2	Fe^{2+}/Fe	Zn^{2+}/Zn	Li^+/Li
ε^0/V	1.35827	0.7996	0.771	0.521	0.3419	0.22233	0.00000	−0.447	−0.7618	−3.0401

Using (1.188) and (1.237), we can write

$$\Delta G^0 = -nF\Delta\varepsilon^0 \tag{7.113}$$

Taking into consideration the zero of standard electrode potentials (the hydrogen standard electrode potential), (7.113) can be replaced by

$$G^0 = -nF\varepsilon^0 \tag{7.114}$$

Let us now consider a metal M with oxidation states M^{m+} and M^{n+} ($n > m$). Being an extensive thermodynamic function, the Gibbs energy is additive, and so we can write

$$G^0_{n0} = G^0_{nm} + G^0_{m0} \tag{7.115}$$

where subscripts $n0$, nm, and $m0$ express the reductions $n \to 0$, $n \to m$, $m \to 0$, respectively. Dividing both members of (7.115) by $-F$ and using (7.114), we obtain

$$n\varepsilon_{n0}^{0} = (n-m)\varepsilon_{nm}^{0} + m\varepsilon_{m0}^{0} \tag{7.116}$$

This result allows us to determine any one of these standard potentials, provided we know the other two. For instance, if (7.116) is applied to the standard potentials of iron in acidic solution, we can write

$$3\varepsilon_{30}^{0} = \varepsilon_{32}^{0} + 2\varepsilon_{20}^{0} \tag{7.117}$$

(**E14**, **E15**).

Let us consider now manganese and its various oxidation states in acidic solution: Mn(VII) as MnO_4^-, Mn(VI) as MnO_4^{2-}, Mn(IV) as MnO_2, Mn^{3+}, Mn^{2+}, and Mn. The **Latimer diagram** for this element shows the standard electrode potentials connecting the various oxidation states of the element (Fig. 7.46).

A graph of $n\varepsilon^0$ as a function of n is called **Frost diagram**, where each point corresponds to the pair of values (n, $n\varepsilon^0$) and the slope between two consecutive points m and n is given by

$$(n-m)\varepsilon_{nm}^{0}/(n-m) \tag{7.118}$$

Figure 7.47 shows a Frost diagram for manganese, in acidic aqueous solution. This diagram can be obtained using the standard electrode potentials of the Latimer diagram of Fig. 7.46, with the values to be plotted on the y-axis being obtained by multiplying the number of electrons involved in the reduction half-reaction by the corresponding standard electrode potential.

By the Latimer diagram, it can be seen that the reduction pair Mn^{3+}/Mn^{2+} has a standard reduction potential, $\varepsilon^0 = 1.5415$ V, greater than that of the pair Mn^{2+}/Mn, $\varepsilon^0 = -1.185$ V. Therefore, the oxidation–reduction reaction involving Mn^{3+}, Mn^{2+}, and Mn in acidic solution is

$$2Mn^{3+} + Mn \to 3Mn^{2+} \tag{7.119}$$

An oxidation-reduction reaction like (7.119) in which two oxidation states of the same element transform into the intermediate oxidation state is called a **compro-portionation reaction**, and the Frost diagram is concave at the intermediate oxidation

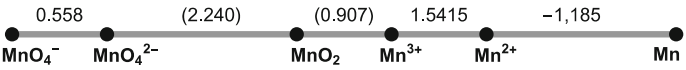


Fig. 7.46 Latimer diagram for manganese with its various oxidation states in acidic aqueous solution. The standard electrode potentials are expressed in volts (V) and were extracted from Haynes (2011), with those inside parentheses having been calculated. This figure was obtained with *Mathematica*

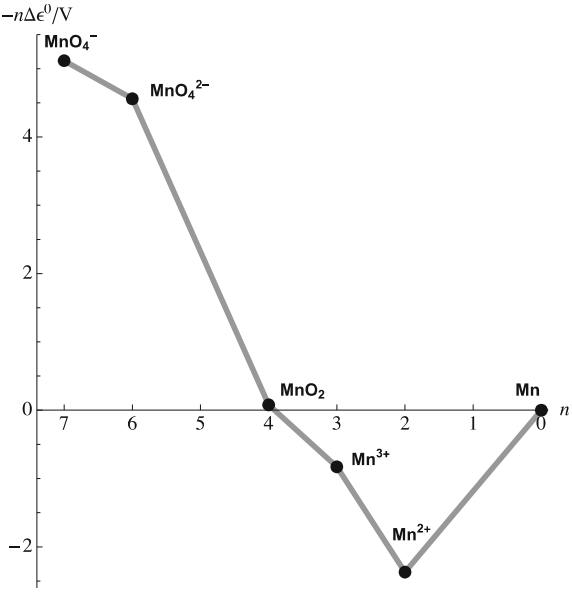


Fig. 7.47 Frost diagram for manganese in acidic aqueous solution. This figure was obtained with *Mathematica*

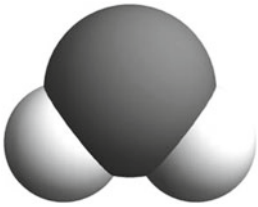
state (Fig. 7.47). In turn, an oxidation–reduction reaction in which one intermediate oxidation state of a chemical element transforms into two oxidation states of the same element is called a **disproportionation reaction**, and the corresponding Frost diagram is convex at the intermediate oxidation state (**E16**, **E17**, **E18**).

Mathematica Codes

M1. Representation of the Water Molecule

The following *Mathematica* code represents the H_2O molecule by an intersection of spheres using the command `Sphere[p,r]`, which represents a sphere of radius `r` centered at the point `p`:

```
(* The following Cartesian coordinates for atoms in H2O are consistent with ROH=0.9611 A
and <HOH=104.5296 *)
(* It is assumed that the ratio of atomic radii H/O = 0.7 *)
Graphics3D[{White,Sphere[{0,0.760117,-0.470585},0.7],
   Sphere[{0.0,-0.760117,-0.470585},0.7],Gray,Sphere[{0, 0, 0.117646}]},
   Lighting→"Neutral",Boxed→False,ViewPoint→Right]
```



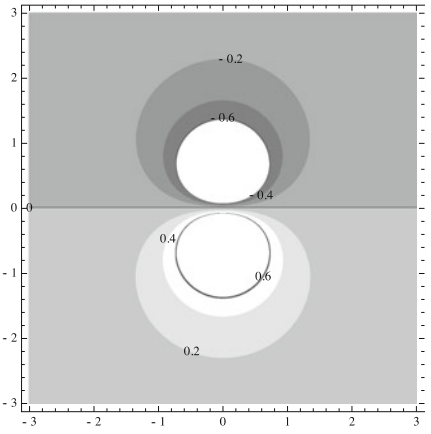
Suggestion: Modify the code to give specularity to the H_2O representation. To this end, consult Wolfram Documentation in the Help Section of *Mathematica*. Try different specular exponents.

M2. Electrostatic Potential Contours for a Dipole

The following *Mathematica* code generates contour plots in the *xy*-plane for the electrostatic potential of a distribution of two charges (−1 and 1), located on the *y*-axis, at $y = 0.5$ and $y = -0.5$. The code begins by defining the function `f[q,p, r]`, which gives the electrostatic potential of a static distribution of electric charges [see (5.13)]. The *Mathematica* command `Lighter[Gray,i]` gives the fraction `i` of `Gray` (`i=0` corresponds to `Gray` and `i=1` to `White`), and the *Mathematica* command `ContourPlot` generates the contours of a function in the *xy*-plane, which is the electrostatic potential of a dipole:

```
Clear[f]
f[q_,p_,r_] := Sum[q[[i]]/Norm[r-p[[i]]],{i,Length[q]}]
q={-1,1}
Length[q]
q[[1]]
q[[2]]
p={{0.,0.5},{0.,-0.5}}
p[[1]]
p[[2]]
c=Table[Lighter[Gray,i],{i,0,1,1/6}]
ContourPlot[f[q,p,{x,y}],{x,-3,3},{y,-3,3},
  ContourShading→c,ContourLabels→True,PlotPoints→150,ImageSize→{300,300}]
```

```
{-1, 1}

2

-1

1

{{0., 0.5}, {0., -0.5}}

{0., 0.5}

{0., -0.5}
```

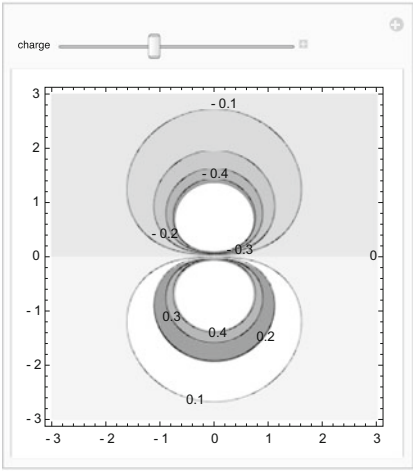


Suggestion: Try `Table[Lighter[Gray,i],{i,0,1,1/6}]` with other colors.

M3. Interactive Manipulation of Charge

This *Mathematica* code uses `Manipulate` to control the absolute value of the charge in the electrostatic potential of a dipole:

```
Clear[f]
f[q_,p_,r_]:=Sum[q[[i]]/Norm[r-p[[i]]],{i,Length[q]}]
p={{0.,0.5},{0.,-0.5}};
c=Table[Lighter[Gray,i],{i,0,1,1/6}];
Manipulate[ContourPlot[f[{-a,a},p,{x,y}],{x,-3,3},{y,-3,3},
  ContourShading→c,ContourLabels→True,PlotPoints→150,ImageSize→{300,300}],
  {{a,0.70,"charge"},0.50,1.00,0.02}]
```



Suggestion: Run the code and observe the contour changes as the absolute value of the charge is increased.

M4. Vector Field Streamlines

We begin by considering a dipole with a charge -1 at $(x, y) = (-1, 0)$ and charge 1 at $(x, y) = (1, 0)$. For this dipole, the electrostatic potential in the xy-plane is given by

$$\phi(x, y) = \frac{1}{r_+} - \frac{1}{r_-}$$

where

$$r_+ = \left[(x-1)^2+y^2\right]^{1/2} \quad r_- = \left[(x+1)^2+y^2\right]^{1/2}$$

Substitution of these equalities in the electrostatic potential gives

$$\phi(x,y) = \frac{1}{r_+} - \frac{1}{r_-} = \frac{1}{\left[(x-1)^2+y^2\right]^{1/2}} - \frac{1}{\left[(x+1)^2+y^2\right]^{1/2}}$$

The x- and y-components of the electric field are given by

$$\mathcal{E}_x = -\frac{\partial\phi}{\partial x} = \frac{x-1}{\left[(x-1)^2+y^2\right]^{3/2}} - \frac{x+1}{\left[(x+1)^2+y^2\right]^{3/2}}$$
$$\mathcal{E}_y = -\frac{\partial\phi}{\partial y} = \frac{y}{\left[(x-1)^2+y^2\right]^{3/2}} - \frac{y}{\left[(x+1)^2+y^2\right]^{3/2}}$$

For a point dipole with charges -1 and 1 separated by an infinitesimal distance δ, we have

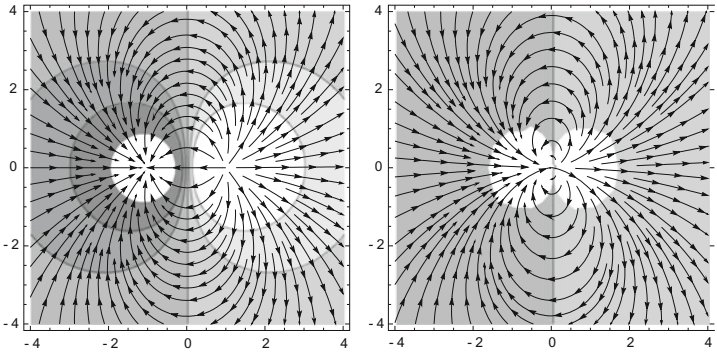
$$\phi(\delta) = \frac{\delta x}{r^3} = \frac{\delta x}{(x^2+y^2)^{3/2}}$$

and

$$\mathcal{E}_x = -\frac{\partial\phi}{\partial x} = \frac{3\delta x^2}{(x^2+y^2)^{5/2}} - \frac{\delta}{(x^2+y^2)^{3/2}} \quad \mathcal{E}_y = -\frac{\partial\phi}{\partial y} = \frac{3\delta xy}{(x^2+y^2)^{5/2}}$$

The above equations are used in the following *Mathematica* code, where we have set $\delta = 0.01$. The command `StreamPlot` plots the vector field as a function of x and y. Observe the difference between the contours for the dipole and the point dipole:

```
Clear[f]
f[q_,p_,r_]:=Sum[q[[i]]/Norm[r-p[[i]]],{i,Length[q]}]
q={-1,1};
p={{-1,0},{1,0}};
d=0.01; (* = distance between -q and +q, for the point dipole *)
c=Table[Lighter[Gray,i],{i,0,1,1/6}];
dipole=ContourPlot[Evaluate[f[q,p,{x,y}]],{x,-4,4},{y,-4,4},
    Contours→{-0.75,-0.25,-0.1,0,0.1,0.25,0.75},ContourShading→c];
pointdipole=ContourPlot[(d x)/(x^2+y^2)^(3/2),{x,-4,4},{y,-4,4},
    Contours→{-0.75,-0.25,-0.1,0,0.1,0.25,0.75},ContourShading→c];
vectorlinesdipole=StreamPlot[{(-1+x)/(((-1+x)^2+y^2)^(3/2)+0.0001)-
    (1+x)/(((1+x)^2+y^2)^(3/2)+0.0001),y/(((-1+x)^2+y^2)^(3/2)+0.0001)-
    y/(((1+x)^2+y^2)^(3/2)+0.0001)},{x,-4,4},{y,-4,4},StreamStyle→Black];
vectorlinespointdipole=StreamPlot[{(3d x^2)/(x^2+y^2)^(5/2)-d/(x^2+y^2)^(3/2),
    (3d x y)/(x^2+y^2)^(5/2)},{x,-4,4},{y,-4,4},StreamStyle→Black];
aa=Show[{dipole,vectorlinesdipole},ImageSize→{300,300}];
bb=Show[{pointdipole,vectorlinespointdipole},ImageSize→{300,300}];
Row[{aa, bb}, "      "]
```



Suggestion: Run the above code with different values for `d`, for example 0.1 and 0.5, and observe the differences.

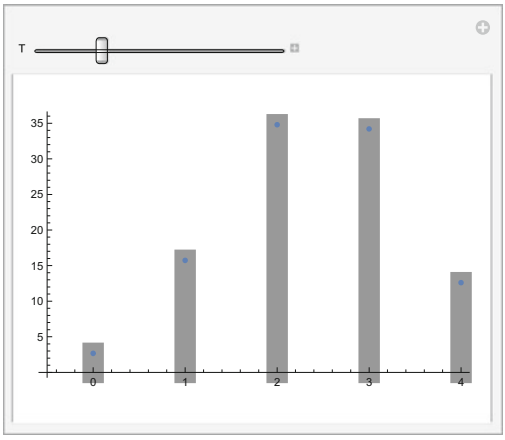
M5. Luzar's Model

Assume that in liquid water, there are five types of H_2O molecules according to the number of hydrogen bonds each molecule can form, namely, 0, 1, 2, 3, or 4. This *Mathematica* code uses (7.26) to calculate the distribution of hydrogen bonds in water according to Luzar's model as temperature varies between 273 and 373 K, in steps of 1 K. The *Mathematica* function `Binomial[n,k]` gives the binomial coefficient $\binom{n}{k}$, and the following line of code calculates $\binom{4}{k}$ for $k = 0, 1, 2, 3, 4$:

```
Table[Binomial[4,k],{k,0,4}]
```

```
{1, 4, 6, 4, 1}
```

```
R = 0.00831446;      (* kJ mol-1 K-1 *)
α = 178.; ε = 13.8; (* kJ mol-1 *)
Manipulate[ListPlot[Table[{k,100 Binomial[4,k]*(α*Exp[-ε/(R*T)])^(4-k)
    /(1+α*Exp[-ε/(R*T)])^4},{k,0,4}],AxesOrigin→{-0.5,0},Filling→Axis,
    FillingStyle→Directive[Gray,Thickness[0.05]],ImageSize→{350,250}],
    {{T,298,"T"},273,373,1}]
```



Suggestion: Run the above *Mathematica* code and find how the hydrogen bond network changes with temperature.

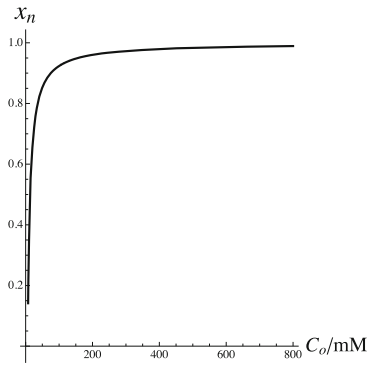
M6. Micelle–Monomers Equilibrium

The following *Mathematica* code shows how to solve (7.71) by the iteration process given by (7.72). The independent variable represented in the code by `x` is the total concentration of monomers C_0, and # is the mole fraction of monomers in the micelles x_n that appears in both members of (7.72).

In the code below, the *Mathematica* function `Nest` applies three times (three iterations) to `1-(x*10^-3)^(1./n-1)*(#/(n*10^(2 n)))^(1/n)&`. In this expression [see (7.71)], x in millimolar units is multiplied by 10^{-3} to convert into molar units, and the equilibrium constant K_n is made equal to 10^{2n}. The result of the third iteration is given by `iteration[[3]]`:

```
n = 20;
iteration=Table[Nest[1-(x*10^-3)^(1./n-1)*(#/(n 10^(2 n)))^(1/n)&,x,i],{i,1,3}];
iteration[[3]]
Plot[iteration[[3]],{x,0,800},AxesOrigin→{0.,0.},AxesLabel→{Style["C_o/mM",Large,Black],
    Style["x_n",Large,Black]},PlotStyle→{Black,Thickness[0.008]},AspectRatio→1.2,
    ImageSize→{350,350}]
```

$$1-\frac{6.09465\left(1-\frac{6.09465\left(1-\frac{6.09465}{x^{0.95}}\right)^{1/20}}{x^{0.95}}\right)^{1/20}}{x^{0.95}}$$

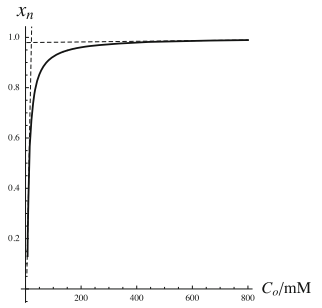


Suggestion: use a pocket calculator to verify `6.09465` in the result for `iteration[[3]]`.

M7. Critical Micelle Concentration

The following *Mathematica* code shows how to obtain an estimate for the critical micelle concentration (CMC $\approx$ 21 mM) through the intersection of tangents to the x_n-versus-C_0 curve, at C_0 = 10 mM and C_0 = 800 mM. The intersection of these tangents is the point at which `h[x]` and `g[x]` (the derivatives of `f[x]`) are equal. In order to solve the equation that expresses this equality, the code uses `FindRoot` to search a numerical root starting from `x=50`:

```
Clear[f,g,h]
n=20;
f[x_]:=Nest[1 - (x*10^-3)^(1/n-1) (#/(n 10^(2 n)))^(1/n)&,x,3];
g[x_]:=f[10.]+(x-10.)*(f[10.00005]-f[9.99995])/0.0001
h[x_]:=f[800.]+(x-800.)*(f[800.00005]-f[799.99995])/0.0001
aa=Plot[f[x],{x,5.,800.0},AxesOrigin→{0.,0.},AxesLabel→{Style["C0/mM",Large,Black],
    Style["xn",Large,Black]},PlotStyle→{Black,Thickness[0.008]},AspectRatio→1.2];
bb=Plot[{h[x],g[x]},{x,5.,800.0},PlotStyle→{{Black,Dashed,Thickness[0.003]},
    {Black,Dashed,Thickness[0.003]}}];
graph=Show[aa,bb]
cmc = SetPrecision[FindRoot[h[x]-g[x]==0,{x,50}][[1]],2]
SetPrecision[x/.cmc,2]
```

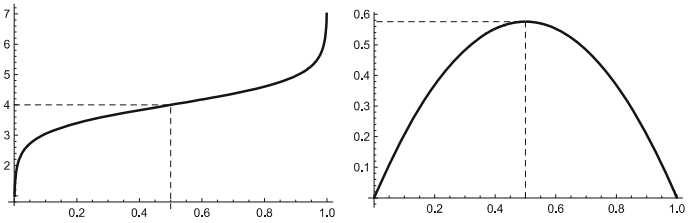



```
21.
```

M8. Weak Acid HA

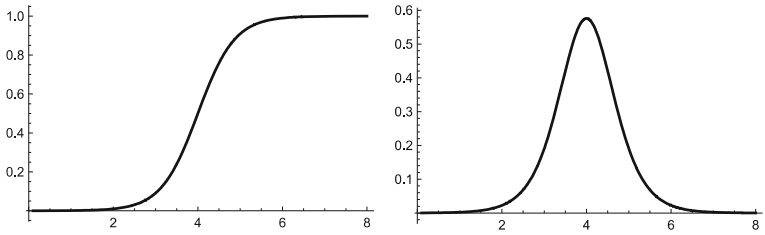
The way in which the solution of a weak acid HA resists a change of pH can be determined by evaluating dx_{A^-}/dpH. The following *Mathematica* code considers the pH as a function of x_{A^-} [see (7.96)] for $pK_{HA} = 4$, and evaluates dx_{A^-}/dpH by taking the inverse of the derivative of $pH = f(x_{A^-})$, as a function of x_{A^-}, at right. Note that the range of x_{A^-} values is [0.001, 0.999], since 0 and 1 would give rise to singularities. It is shown that the solution attains maximum buffer capacity at $x_{A^-} = 0.5$, when HA and A^- have equal mole fractions. The code makes use of `Line` for representing the dashed line segments joining a sequence of points:

```
 (* x = x(A-); f[x] = pH; pKa = 4 *)
f[x_]:=4.-Log[10,(1-x)/x]
max = FindMaximum[1/f'[x],{x,0.5}];
aa=Plot[f[x],{x,0.001,0.999},PlotStyle→{Black,Thickness[0.008]},ImageSize→{250,250}];
bb=Plot[1/f'[x],{x,0.001,0.999},PlotStyle→{Black,Thickness[0.008]},ImageSize→{250,250}];
cc= Graphics[{Dashed,Line[{{0,4},{0.5,4},{0.5,0.}}]}];
dd= Graphics[{Dashed,Line[{{0.5,0.},{0.5,max[[1]]},{0.,max[[1]]}}]}];
rr=Show[aa,cc];
ss=Show[bb,dd];
Row[{rr,ss},"    "]
```



The following *Mathematica* code plots x_{A^-} and $dx_{A^-}/d\text{pH}$ as functions of pH [see (7.97)] for $pK_{HA} = 4$:

```
(* x = pH; f[x] = xA- *)
f[x_]:=1/(1+10^(4-x))
aa=Plot[f[x],{x,0,8},PlotStyle→{Black,Thickness[0.008]},ImageSize→{350,350}];
bb=Plot[f'[x],{x,0,8},PlotStyle→{Black,Thickness[0.008]},ImageSize→{350,350}];
Row[{aa,bb},"   "]
```



Glossary

Bronsted–Lowry acid — A proton (H^+) donor. A Bronsted–Lowry base is a proton acceptor.

Buffer solution — A solution that resists pH changes and, for an aqueous solution of weak acid HA, attains maximum buffer capacity at $x_{A^-} = 0.5$, when HA and A^- have the same mole fractions.

Comproportionation — An oxidation–reduction reaction whereby two chemical species with different oxidation numbers in one element form a product in which that same element has an intermediate oxidation number.

Critical micelle concentration — The concentration of a surfactant above which micelles form.

Debye–Hückel limiting law — Enables one to determine the activity coefficient of an electrolyte at very low concentrations in water as a function of the ionic strength [see (7.63)], a concept that was introduced in 1921 by Lewis (1875–1946) and Randall (1888–1950) and is a measure of the concentration of ions in a solution [see (7.64)].

Debye–Hückel theory — Was formulated in 1923 by Peter Debye (1884–1966) and Eric Hückel (1896–1980). It evaluates departures from ideal behavior in dilute ionic solutions

Debye length	The distance over which a charge in an ionic solution is shielded by the charge distribution of ions of opposite charge.
Disproportionation	An oxidation–reduction reaction in which a reactant with one element in a particular oxidation state is simultaneously oxidized and reduced, giving two different products.
Lewis acid	An electrophile that accepts an electron pair from a Lewis base to form a Lewis adduct.
Micelle	An aggregate of amphipathic molecules with colloidal dimensions (from 1 to 1000 nm) that in aqueous solution has the hydrophilic head of the amphipathic molecules exposed to the solvent and the hydrophobic tails sequestered in the micelle center.
Ostwald coefficient	The ratio of the gas volume change ΔV_g due to the irreversible dissolution of the gas in water to the volume of the aqueous solution V_{sln}, at equilibrium [see (7.36) and Fig. 7.33].
Quadrupole	Charge distribution with null total charge and null total dipole that consists of four charges of equal magnitude, two positive and two negative.
Surfactant	Amphipathic molecule that decreases the surface tension of an aqueous mixture with respect to the surface tension of pure water (71.99 mJ m^{-2} at 25 °C). In the strict technical sense, a surfactant is a compound that for a concentration not greater than 0.01 M, lowers the water surface tension by 30 mJ m^{-2} or more.

Exercises

E1. Considering the dipole moment definition, for a static distribution of point charges, show that the dipole moment vector points to the positive charge.
E2. Show that the dipole moment of a charge distribution whose total charge is zero is independent of the origin.
E3. Considering the electrostatic potential atomic charges of H_2O, $q_O = -0.680$ and $q_H = 0.340$, and the geometric values $R_{OH} = 0.9611$ Å and $\alpha_{HOH} = 104.53°$, obtain the molecular dipole moment.

E4. Use *Mathematica* to plot the angular functions involved in the δ and δ^3 terms of (7.13). Show that those functions are proportional to the $Y_1^0(\theta,0)$ and $Y_3^0(\theta,0)$ spherical harmonics in the plane.
E5. Use *Mathematica* to plot the contours of the electrostatic potential of a quadrupole with the charges (−0.225, −0.225, 0.225, 0.225) at the *x*,*y*-coordinates (0., 0.1176), (0., −1.05922), (−1.0, −0.48), and (1., −0.48), respectively. Compare these contours with those of the quadrupole term for H_2O (Fig. 7.7).
E6. At CCCBDB (cccbdb.nist.gov), find the energy of H_2O at equilibrium (no zero-point correction) given by calculations obtained with the B3LYP method and the STO-3G, 3-21G, 6-31G, 6-311G*, 6-311G**, 6-311+G(3df,2p), 6-311+G(3df,2pd), cc-pVQZ, aug-cc-pVQZ basis sets. Use *Mathematica* to plot these data with the energy in eV. Comment on the obtained graph.
E7. Derive (7.22).
E8. Specify the symmetry species of the normal modes of H_2O and DHO.
E9. Chlorine has two natural isotopes, ^{35}Cl and ^{37}Cl, whose natural abundances are in the approximate ratio of 3:1. The breathing mode (symmetric stretching of the carbon tetrachloride molecule can be observed in the Raman spectrum of this liquid. Give an estimate of the relative Raman intensities of the distinct isotopic lines of the breathing mode of the carbon tetrachloride molecule and the frequency difference in cm^{-1} between the lines corresponding to the isotopic combinations (35, 35, 35, 35) and (35, 35, 35, 37). The breathing vibration of the (35, 35, 35, 35) isotopic combination occurs at approximately 459 cm^{-1} (CCCBDB).
E10. Consider p_{hb} as a function of ε, for T = 298.15 K, and assume α = 178 kJ mol^{-1}. Use *Mathematica* to plot the function $p_{hb}(\varepsilon)$ and its first derivative with respect to ε, and determine the inflection point of $p_{hb}(\varepsilon)$.
E11. Consider a dilute solution of one particular nonpolar solute and convert mole fraction into molarity. What is the value of the conversion factor when the solvent is water?
E12. Calculate the mean activity coefficient for a 0.001 molal aqueous solution of $CaCl_2$.
E13. An oxidizing agent that converts H_2O into O_2 and a reducing agent that reduces H_2O to H_2 are not thermodynamically stable in aqueous solution. Derive the potentials of the O_2/H_2O and H_2O/H_2 reductions as functions of pH at constant temperature (T = 298.15 K). Assume that $pO_2 = pH_2 = p^0 = 1$ atm. Making use of *Mathematica*, plot the obtained functions and fill with gray the region of thermodynamic stability for water.
E14. Consider MnO_2, Mn^{3+}, and Mn^{2+} in acidic aqueous solution. Evaluate $\varepsilon^0(MnO_2/Mn^{3+})$, knowing that $\varepsilon^0(MnO_2/Mn^{2+})$ = 1.224 V and $\varepsilon^0(Mn^{3+}/Mn^{2+})$ = 1.5415 V.
E15. Consider MnO_4^-, MnO_4^{2-}, and MnO_2 in acidic aqueous solution. Evaluate $\varepsilon^0(MnO_4^{2-}/MnO_2)$, knowing that $\varepsilon^0(MnO_4^-/MnO_4^{2-})$ = 0.558 V and $\varepsilon^0(MnO_4^-/MnO_2)$ = 1.679 V.
E16. Consider $\varepsilon^0(Fe^{3+}/Fe^{2+})$ = 0.771 V and $\varepsilon^0(Fe^{2+}/Fe)$ = -0.447 V and an acidic aqueous solution consisting of Fe(III), Fe(II), and Fe(0). Use

Mathematica to obtain the Frost diagram for iron in acidic aqueous solution. How are Fe^{3+}, Fe^{2+}, and Fe(s) related, by disproportionation or comproportionation? Explain.

E17. Consider $\varepsilon^0(Cu^{2+}/Cu^+) = 0.153$ V and $\varepsilon^0(Cu^+/Cu) = 0.521$ V and an acidic aqueous solution consisting of Cu(II), Cu(I), and Cu(0). Use *Mathematica* to obtain the Frost diagram for copper. How are Cu^{2+}, Cu^+, and Cu(s) related, by disproportionation or comproportionation? Explain.

E18. Consider the electrochemical concentration cell $Ag|Ag^+||Ag^+|Ag$, and represent by a_a and a_c the silver ion activities at the anode and cathode, respectively. In which electrode, anode or cathode, is the activity of the silver ion greater? Why? Write the Nernst equation for this cell.

References

Martin JML, de Oliveira G (1999) J Chem Phys 111:1843

Parthiban S, Martin JML (2001) J Chem Phys 114:6014

Aziz RA, Slaman MJ (1991) An examination of ab initio results for the helium potential energy curve. J Chem Phys 94:8047–8053

Woon DE (1991) Benchmark calculations with correlated molecular wave functions. V. The determination of accurate ab initio intermolecular potentials for He_2, Ne_2, and Ar_2. J Chem Phys 100:2838–2850

Seuss E, Bohrmann G, Greinert J, Lausch E (1999) Flammable ice. Sci Am

Luzar A, Svetina S, Zeks B (1983) The contribution of hydrogen bonds to the surface tension of water. Chem Phys Lett 96:485–490

Teixeira J, Luzar A (1999) Physics of liquid water. Structure and dynamics. In: Bellissent-Funel MC (ed) Hydration processes in biology, NATO ASI series, Series A: Life sciences, vol 305. Kluwer, Dordrecht, pp 35–65

Haynes WM (Editor-in-Chief) (2011) Handbook of chemistry and physics on CD-ROM, version 2011. CRC Press, Taylor and Francis Group

Ben-Naim A, Baer S (1963) Method for measuring solubilities of slightly soluble gases in liquids. Trans. Faraday Soc. 59:2735–2738

Yaacobi M, Ben-Naim A (1973) Hydrophobic interaction in water-ethanol mixtures. J Solution Chem 2:425–443

Vargaftik NB, Volkov BN, Voljak LD (1983) International tables of the surface tension of water. J Phys Chem Ref Data 12(3):817–820

Berg JC (2010) An introduction to interfaces and colloids—the bridge to nanoscience. World Scientific Publishing Co. Pte. Ltd

Oh SG, Shah DO (1993) The effect of micellar lifetime on the rate of solubilization and detergency in sodium dodecyl sulfate solutions. J Amer Oil Chemists Soc 70:673–678

Andrade-Dias C, Lima S, Teixeira-Dias JJC (2007) From simple amphiphilic to surfactant behavior: analysis of ^1H NMR chemical shift variations. J Colloid Interface Sci 316:31–36

Kapustinskii AF (1956) Lattice energy of ionic crystals. Q Rev Chem Soc 10:283–294

Rini M, Magnes B-Z, Pines E, Nibbering ETJ (2003) Real-time observation of bimodal proton transfer in acid-base pairs in water. Science 301:349–352

Mohammed OF, Pines D, Dreyer J, Pines E, Nibbering ETJ (2005) Sequential proton transfer through water bridges in acid-base reactions. Science 310:83–86

Bandura AV, Lvov SN (2006) The ionization constant of water over wide ranges of temperature and density. J Phys Chem Ref Data 35(1):15–30

Covington AK (2011) pH scale for aqueous solutions. In: Haynes WM (Editor-in-Chief) Handbook of chemistry and physics on CD-ROM, version 2011. CRC Press, Taylor and Francis Group

Further Reading

Levine IN (2013) Quantum chemistry, 7th edn. Pearson Education Inc.
Lynden-Bell RM, Morris SC, Barrow JD, Finney JL, Harper CL Jr (2010) Water and life: the unique properties of H_2O. CRC Press, Taylor & Francis Group
Atkins P, Paula J (2006) Physical chemistry for the life sciences. Oxford University Press and W.H. Freeman and Company
Kondepudi D (2008) Introduction to modern thermodynamics. Wiley, New York

Appendix

Physical Constants

Constant	Symbol	SI value
Speed of light in vacuum	c	299,792,458 m s^{-1}
Elementary charge	e	$1.602176565 \times 10^{-19}$ C
Permittivity of vacuum	ε_0	$8.854187817 \times 10^{-12}$ C V^{-1} m^{-1}
Planck constant	h	$6.62606957 \times 10^{-34}$ J s
Electron mass	m_e	$9.10938291 \times 10^{-31}$ kg
Proton mass	m_p	$1.672621777 \times 10^{-27}$ kg
Neutron mass	m_n	$1.674927351 \times 10^{-27}$ kg
Deuteron mass	m_d	$3.34358348 \times 10^{-27}$ kg
Avogadro constant	N_A	$6.02214129 \times 10^{23}$ mol^{-1}
Faraday constant	$F = eN_A$	96,485.3365 C mol^{-1}
Boltzmann constant	k_B	$1.3806488 \times 10^{-23}$ J K^{-1}
Gas constant	$R = k_B N_A$	8.3144621 J mol^{-1} K^{-1}
Atomic mass constant	m_u	$1.660538921 \times 10^{-27}$ kg
Hartree energy	E_h	$4.35974434 \times 10^{-18}$ J
Bohr radius	$a_0 = \varepsilon_0 h^2/(\pi m_e e^2)$	$0.52917721092 \times 10^{-10}$ m
Electron volt	eV	$1.602176565 \times 10^{-19}$ J

P.J. Mohr, B.N. Taylor, and D.B. Newell (2011), "The 2010 CODATA Recommended Values of the Fundamental Physical Constants" (Web Version 6.0). This database was developed by J. Baker, M. Douma, and S. Kotochigova. Available: http://physics.nist.gov/constants. National Institute of Standards and Technology, Gaithersburg, MD 20899. Alternatively, each of these physical constants can be obtained from the computational knowledge engine Wolfram Alpha at www.woframalpha.com.

J.J.C. Teixeira-Dias, *Molecular Physical Chemistry*,
DOI 10.1007/978-3-319-41093-7

Conversion Factors

1 hartree = 27.211385 eV = 2625.4996 kJ mol^{-1}
1 eV = 96.4853365 kJ mol^{-1}
1 cm^{-1} = 0.011962657 kJ mol^{-1}
1 cal = 4.184 J
1 eÅ = 4.8032 D

Atomic Units

Action	$h/(2\pi)$	$1.054571726 \times 10^{-34}$ J s
Charge	e	$1.602176565 \times 10^{-19}$ C
Length	a_0	$0.52917721092 \times 10^{-10}$ m
Mass	m_e	$9.10938291 \times 10^{-31}$ kg
Electric constant	$4\pi\varepsilon_0$	$1.112650056 \times 10^{-10}$ C V^{-1} m^{-1}

Frequently Used Prefixes

Factor	Name	Symbol	Factor	Name	Symbol
10^{18}	exa	E	10^{-3}	milli	m
10^{15}	peta	P	10^{-6}	micro	μ
10^{12}	tera	T	10^{-9}	nano	n
10^{9}	giga	G	10^{-12}	pico	p
10^{6}	mega	M	10^{-15}	femto	f
10^{3}	kilo	k	10^{-18}	atto	a

Greek Alphabet

Alpha	A	α	Iota	I	ι	Rho	P	ρ
Beta	B	β	Kappa	K	κ	Sigma	Σ	σ
Gamma	Γ	γ	Lambda	Λ	λ	Tau	T	τ
Delta	Δ	δ	Mu	M	μ	Upsilon	ϒ	υ
Epsilon	E	ε	Nu	N	ν	Phi	Φ	ϕ
Zeta	Z	ζ	Xi	Ξ	ξ	Chi	X	χ
Eta	H	η	Omicron	O	o	Psi	Ψ	ψ
Theta	Θ	θ	Pi	Π	π	Omega	Ω	ω

Integrals

Mathematica instruction	Output
`Integrate[1/x,x]`	`Log[x]`
`Integrate[x^n,x]`	`x^(1+n)/(1+n)`
`Integrate[Sin[x],x]`	`-Cos[x]`
`Integrate[Cos[x],x]`	`Sin[x]`
`Integrate[Tan[x],x]`	`-Log[Cos[x]]`
`Integrate[Exp[-x^2],x]`	`1/2 Sqrt[π] Erf[x]`
`Integrate[Exp[-b x^2],{x,0,Infinity}, Assumptions→b>0]`	`Sqrt[π]/(2 Sqrt[b])`

Extensive mathematical data can be obtained from Wolfram Alpha at www.woframalpha.com.

Character Tables of Point Groups

Available at http://symmetry.jacobs-university.de

Answers to Exercises

Chapter 1—Thermodynamics

E1. *Mathematica* Code and Results

```
1/Integrate[Exp[-m(vx^2+vy^2+vz^2)/(2k T)],{vx,-Infinity,Infinity},
    {vy,-Infinity,Infinity},{vz,-Infinity,Infinity},Assumptions→{m>0,k>0,T>0}]
```

$$\frac{1}{2\sqrt{2}\,\pi^{3/2}\left(\frac{k\,T}{m}\right)^{3/2}}$$

E2. *Mathematica* Code and Results

```
Clear[k,T,m]
norm=1/Integrate[Exp[-m(vx^2+vy^2+vz^2)/(2k T)],{vx,-Infinity,Infinity},
  {vy,-Infinity,Infinity},{vz,-Infinity,Infinity},Assumptions→{m>0,k>0,T>0}];
f[v_,T_]:=4Pi*norm*Exp[-m*v^2/(2k T)]*v^2
Print["Most probable speed"]
sol=Solve[D[f[v,T],v]==0,v][[3]];
Print["Mean speed"]
vm=Integrate[v*f[v,T],{v,0,Infinity},Assumptions→{m>0,k>0,T>0}]
Print["Root mean square speed"]
vrms=Sqrt[Integrate[f[v,T]*v^2,{v,0,Infinity},Assumptions→{m>0,k>0,T>0}]]
k=1.3806488 10^-23; T=298.15; m=0.032/(6.02214129 10^23);
Print["Most probable speed for O2 at 298.15 K"]
SetPrecision[v/.sol,3]
Print["Mean speed for O2 at 298.15 K"]
SetPrecision[vm,3]
Print["Root mean square speed for O2 at 298.15 K"]
SetPrecision[vrms,3]
```


J.J.C. Teixeira-Dias, *Molecular Physical Chemistry*,
DOI 10.1007/978-3-319-41093-7

```
Most probable speed
```

$$\left\{v \to \frac{\sqrt{2}\ \sqrt{k}\ \sqrt{T}}{\sqrt{m}}\right\}$$

```
Mean speed
```

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

```
Root mean square speed
```

$$\sqrt{3}\sqrt{\frac{k\,T}{m}}$$

```
Most probable speed for O2 at 298.15 K

394.

Mean speed for O2 at 298.15 K

444.

Root mean square speed for O2 at 298.15 K

482.
```

E3. *Mathematica* Code and Results

```
k=1.3806488 10^-23; avogadro=6.02214129 10^23;
p=101325; T=273.15;
d=0.3 10^-9; m=0.032/avogadro;
ro=SetPrecision[p/(k*T),2]; (* molecules/m^3 *)
averagedist=(1/ro)^(1/3); (* m *)
meanfp=SetPrecision[1/(Sqrt[2]*Pi*d^2*ro),3]; (* m *)
rmsspeed=SetPrecision[Sqrt[3k*T/m],3]; (* m s^-1 *)
t1coll=meanfp/rmsspeed; (* time for one collision / s *)
Print["Mean free path / Å"]
meanfp*10^10
Print["Average distance between molecules / Å"]
averagedist*10^10
Print["Average time for one collision / ps"]
t1coll*10^12
 Mean free path / Å

 931.

 Average distance between molecules / Å

 33.

 Average time for one collision / ps

 202.
```

E4. *Mathematica* Code and Results

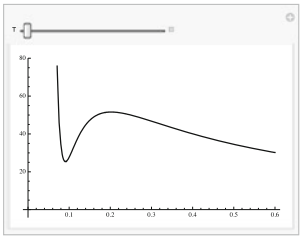
```
f[V_,T_]:=R*T/(V-b)-a/V^2
sol=Solve[D[f[V,T],V]==0&&D[f[V,T],{V,2}]==0,{V,T}][[1]]
p=f[V,T]/.sol
```

$$\left\{V \to 3\,b,\ T \to \frac{8\,a}{27\,b\,R}\right\}$$

$$\frac{a}{27\,b^2}$$

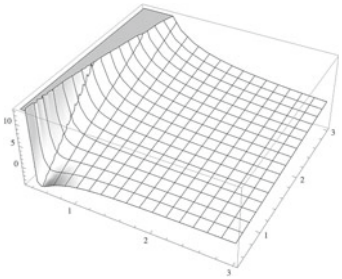
E5. *Mathematica* Code and Results

```
Manipulate[Plot[0.083145*T/(V-0.0429)-3.658/V^2,{V,0.06,0.6},PlotStyle→Black,
AxesOrigin→{0,0}],{T,274,344,5}]
```



E6. *Mathematica* Code and Results

```
p[V_,T_]:=(8/3)*T/(V-1/3)-3/V^2
Plot3D[p[V,T],{V,0.35,3.0},{T,0.5,3.0},PlotSyle→Gray,Lighting→"Neutral"]
```



E7.

$$4x^2y^3\mathrm{d}x + 3x^3y^2\mathrm{d}y = g(x,y)\mathrm{d}x + h(x,y)\mathrm{d}y$$

$$\frac{\partial g(x,y)}{\partial y} = 12x^2y^2 \quad \frac{\partial h(x,y)}{\partial x} = 9x^2y^2$$

$$\frac{\partial g(x,y)}{\partial y} \neq \frac{\partial h(x,y)}{\partial x}$$

Therefore, $4x^2y^3\mathrm{d}x + 3x^3y^2\mathrm{d}y$ is an inexact differential.

E8.

$$p\mathrm{d}V + V\mathrm{d}p = \mathrm{d}(pV) \quad \frac{\partial p}{\partial p} = 1 \quad \frac{\partial V}{\partial V} = 1$$

Therefore, $p\mathrm{d}V + V\mathrm{d}p$ is an exact differential.

E9.

$$\mathrm{d}U = \mathrm{d}Q + \mathrm{d}W$$

$$\mathrm{d}U = \left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T + \left(\frac{\partial U}{\partial V}\right)_T \mathrm{d}V \quad \mathrm{d}W = -p\mathrm{d}V$$

$$\left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T + \left(\frac{\partial U}{\partial V}\right)_T \mathrm{d}V = \mathrm{d}Q - p\mathrm{d}V \quad \mathrm{d}Q = \left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T + \left[p + \left(\frac{\partial U}{\partial V}\right)_T\right]\mathrm{d}V$$

$$\mathrm{d}Q_p = C_p \mathrm{d}T_p \quad C_V = \left(\frac{\partial U}{\partial T}\right)_V \quad \mathrm{d}V_p = \left(\frac{\partial V}{\partial T}\right)_p \mathrm{d}T_p$$

$$\mathrm{d}Q_p = \left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T_p + \left[p + \left(\frac{\partial U}{\partial V}\right)_T\right]\mathrm{d}V_p$$

$$C_p - C_V = \left[p + \left(\frac{\partial U}{\partial V}\right)_T\right]\left(\frac{\partial V}{\partial T}\right)_p$$

E10.

$$\mathrm{d}S = \frac{\mathrm{d}Q}{T} \quad \mathrm{d}S = \frac{C\mathrm{d}T}{T}$$

$$\Delta S = \int_{273}^{373} \frac{C\mathrm{d}T}{T} = C \int_{273}^{373} \frac{\mathrm{d}T}{T} = 4 \times 25.09 \times \ln\frac{373}{273} = 31.3\,\mathrm{J\,K^{-1}}$$

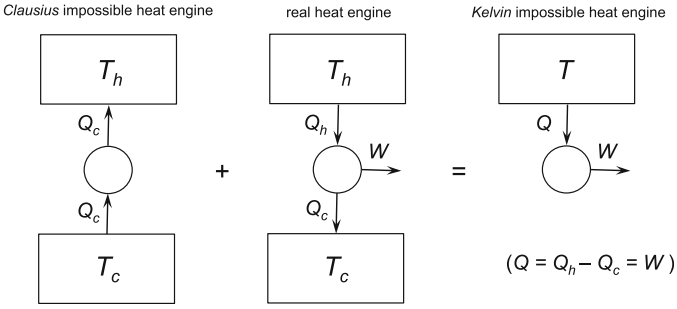
E11.

$$\mathrm{C_3H_{8(g)}} + 5\mathrm{O_{2(g)}} \rightarrow 3\mathrm{CO_{2(g)}} + 4\mathrm{H_2O_{(\ell)}}$$

$$\Delta \mathrm{H}_r^0 = 4\Delta H^0_{f,\mathrm{H_2O_\ell}} + 3\Delta H^0_{f,\mathrm{CO_{2(g)}}} - \Delta H^0_{f,\mathrm{C_3H_{8(g)}}}$$

$$= 4(-285.8) + 3(-393.5) - (-104.7) = -2219\ \mathrm{kJ\ mol^{-1}}$$

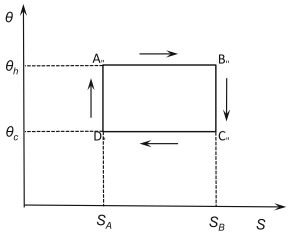
E12.



E13.

$$V = V_0 + V_0 \alpha t$$
$$t = 0\,^{\circ}\mathrm{C} \Rightarrow V = V_0$$
$$\left(\frac{\partial V}{\partial t}\right)_p = V_0 \alpha \Rightarrow \alpha = \frac{1}{V_0}\left(\frac{\partial V}{\partial t}\right)_p$$

E14.



In the temperature–entropy plot of the Carnot cycle, horizontal lines are isotherms (T = constant), and vertical lines are adiabatic transformations ($Q = 0$, that is, S = constant). See **§4** to find out that $\Delta S_{\mathrm{AB}} = Q_{\mathrm{AB}} / \theta_h = -W_{\mathrm{AB}} / \theta_h = R \ln (V_{\mathrm{B}} / V_{\mathrm{A}})$, where $\Delta S_{\mathrm{AB}} > 0$, that is, $S_{\mathrm{B}} > S_{\mathrm{A}}$, because $V_{\mathrm{B}} > V_{\mathrm{A}}$.

E15.

$$
\begin{aligned}
2NO_{2(g)} \rightleftarrows N_2O_{4(g)} \quad \Delta G_r^0 &= \Delta G_{f,N_2O_{4(g)}}^0 - 2\Delta G_{f,NO_{2(g)}}^0 \\
&= 99.8 - 2 \times 51.3 = -2.8\,\mathrm{kJ\,mol^{-1}} \\
\Delta G_r^0 &= RT \ln K_e \\
&- 2.8 \times 10^3 = -8.314462 \times 298.15 \times \ln K_e \quad K_e = 3.09 \\
x_{N_2O_4} = \frac{x}{2-2x+x} &= \frac{x}{2-x} \quad x_{NO_2} = \frac{2-2x}{2-x} \\
K_e = \frac{x_{N_2O_4}}{x_{NO_2}^2} &= \frac{x(2-x)}{(2-2x)^2} = 3.09 \Rightarrow x = 0.7266 \\
x_{N_2O_4} &= 0.5706 \quad x_{NO_2} = 0.4294
\end{aligned}
$$

E16.

$$d\mu_\alpha = -S_{m,\alpha}dT + V_{m,\alpha}dp \quad d\mu_\beta = -S_{m,\beta}dT + V_{m,\beta}dp$$

at equilibrium: $\mu_\alpha = \mu_\beta \quad d\mu_\alpha = d\mu_\beta$

$$
\begin{aligned}
-S_{m,\alpha}dT + V_{m,\alpha}dp = -S_{m,\beta}dT + V_{m,\beta}dp &\Rightarrow \frac{dp}{dT} = \frac{\Delta S_m}{\Delta V_m} = \frac{\Delta H_m}{T\Delta V_m} \\
\frac{dp}{dT} = \frac{\Delta H_m}{T(V_{m,g} - V_{m,\ell})} &\approx \frac{\Delta H_m}{TV_{m,g}} \approx \frac{\Delta H_m p}{RT^2} \Rightarrow \frac{dp}{p} \approx \frac{\Delta H_m}{RT^2}dT
\end{aligned}
$$

E17.

$P = 4(3\,\text{solid phases} + 1\,\text{gas phase})$

$C_{(s)} + O_{2(g)} \rightleftarrows CO_{2(g)} \quad 4Fe_{(s)} + 3O_{2(g)} \rightleftarrows 2Fe_2O_{3(s)}$

$C_i = 5(\text{components}) - 2(\text{chemical equilibria}) = 3$

$P + f = C_i + 2 \quad 4 + f = 3 + 2 \quad f = 1$

variables: $T,\ p_{O_2(g)}, p_{CO_2(g)}, p_t$

equations: $p_t = p_{O_2(g)} + p_{CO_2(g)} \quad p_{CO_2(g)}/p_{O_2(g)} = K_1(T) \quad p_{O_2(g)} = K_2(T)$

Therefore, temperature is the degree of freedom for this system.

E18.

Applying the first and second laws to an isolated system leads to

$$\mathrm{d}U = Td_eS - p\mathrm{d}V \quad \mathrm{d}U = 0,\ \mathrm{d}V = 0,\ d_eS = 0 \quad d_iS \geq 0$$

Therefore, S attains a maximum at the end of the irreversible process, when equilibrium is reached.

E19.

$$X^{n+} + ne^- \rightarrow X \quad \varepsilon = \varepsilon^0 - \frac{RT}{nF}\ln\frac{\mathrm{a}_X}{\mathrm{a}_{X^{n+}}}$$

$$\left(\frac{1}{n}\right)X^{n+} + e^- \rightarrow \left(\frac{1}{n}\right)X \quad \varepsilon = \varepsilon^0 - \frac{RT}{F}\ln\left(\frac{\mathrm{a}_X}{\mathrm{a}_{X^{n+}}}\right)^{1/n} = \varepsilon^0 - \frac{RT}{nF}\ln\frac{\mathrm{a}_X}{\mathrm{a}_{X^{n+}}}$$

E20.

$$2\mathrm{Ag}^+ + \mathrm{Cu} \rightarrow 2\mathrm{Ag} + \mathrm{Cu}^{2+} \quad \Delta\varepsilon^0 = \frac{RT}{nF}\ln K_e \quad \log K_e \approx 16.9\, n\Delta\varepsilon^0$$

$$\mathrm{Ag}^+ + e^- \rightarrow \mathrm{Ag} \quad \varepsilon^0_{\mathrm{Ag}^+/\mathrm{Ag}} = 0.7996\ \mathrm{V}$$

$$\mathrm{Cu}^{2+} + 2e^- \rightarrow \mathrm{Cu} \quad \varepsilon^0_{\mathrm{Cu}^{2+}/\mathrm{Cu}} = 0.3419\ \mathrm{V}$$

$$\Delta\varepsilon^0 = 0.7996 - 0.3419 = 0.4577\ \mathrm{V} \quad K_e \approx 10^{16.9\, n\Delta\varepsilon^0}$$

$$= 10^{16.9\times 2\times 0.4577} = 10^{15.5} = 3.2\times 10^{15}$$

Standard electrode potential values taken from *Handbook of Chemistry and Physics*, 2011.

Chapter 2—Chemical Kinetics

E1. *Mathematica* Code and Results

```
(* [A]/10^-5 M as a function of t/min *)
data = {{0.,5.00},{1.,3.57},{2.,2.49},{3.,1.82},{4.,1.26},
  {5.,0.89},{6.,0.61},{7.,0.45},{8.,0.30},{9.,0.23},{10.,0.16}};
model = a Exp[-k t];
fit=FindFit[data,model,{a,k},t];
f=Function[{t},Evaluate[model/.fit]]
```

```
Function[{t}, 5.01042 e^-0.344655 t]
```

$$[\mathrm{A}] = [\mathrm{A}]_0 \mathrm{e}^{-kt} \quad 0.05[\mathrm{A}]_0 = [\mathrm{A}]_0 \mathrm{e}^{-kt} \quad t = -\frac{\ln 0.05}{k} = -\frac{\ln 0.05}{0.345} = 8.7 \text{ min}$$

E2.

$$a\mathrm{A} + b\mathrm{B} \rightarrow c\mathrm{C} + d\mathrm{D}$$

$$\frac{\mathrm{d}n_\mathrm{A}}{-a} = \frac{\mathrm{d}n_\mathrm{B}}{-b} = \frac{\mathrm{d}n_\mathrm{C}}{c} = \frac{\mathrm{d}n_\mathrm{D}}{d} = \mathrm{d}\xi$$

$$[\mathrm{A}] = n_\mathrm{A}/V \quad [\mathrm{B}] = n_\mathrm{B}/V \quad [\mathrm{C}] = n_\mathrm{C}/V \quad [\mathrm{D}] = n_\mathrm{D}/V$$

$$\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} = \frac{1}{V}\frac{\mathrm{d}n_\mathrm{A}}{\mathrm{d}t} \quad \frac{\mathrm{d}[\mathrm{B}]}{\mathrm{d}t} = \frac{1}{V}\frac{\mathrm{d}n_\mathrm{B}}{\mathrm{d}t} \quad \frac{\mathrm{d}[\mathrm{C}]}{\mathrm{d}t} = \frac{1}{V}\frac{\mathrm{d}n_\mathrm{C}}{\mathrm{d}t} \quad \frac{\mathrm{d}[\mathrm{D}]}{\mathrm{d}t} = \frac{1}{V}\frac{\mathrm{d}n_\mathrm{D}}{\mathrm{d}t}$$

$$\frac{1}{-a}\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} = \frac{1}{-b}\frac{\mathrm{d}[\mathrm{B}]}{\mathrm{d}t} = \frac{1}{c}\frac{\mathrm{d}[\mathrm{C}]}{\mathrm{d}t} = \frac{1}{\mathrm{d}}\frac{\mathrm{d}[\mathrm{D}]}{\mathrm{d}t} = \frac{1}{V}\frac{\mathrm{d}\xi}{\mathrm{d}t}$$

E3. *Mathematica* Code and Results

```
R=0.00831446; (* kJ K^-1 *)
data = {{900,0.01305}, {950,0.07686}, {1000,0.37907}, {1050,1.60593},
  {1100,5.96669},{1150,19.7777}};
fit = SetPrecision[FindFit[data, A Exp[-Ea/(R T)],{A,Ea},T,
  MaxIterations→1000],3]
```

```
{A → 5.57 × 10^12, Ea → 252.}
```

E4.

$$A + \mathrm{B} \rightleftarrows \mathrm{C}$$

$$[\mathrm{A}] = [\mathrm{A}]_{\mathrm{eq}} - \delta \quad [\mathrm{B}] = [\mathrm{B}]_{\mathrm{eq}} - \delta \quad [\mathrm{C}] = [\mathrm{C}]_{\mathrm{eq}} + \delta$$

$$\frac{\mathrm{d}\delta}{\mathrm{d}t} = k_{\rightarrow}[\mathrm{A}][\mathrm{B}] \; - k_{\leftarrow}[\mathrm{C}] = k_{\rightarrow}\left([\mathrm{A}]_{\mathrm{eq}} - \delta\right)\left([\mathrm{B}]_{\mathrm{eq}} - \delta\right) - k_{\leftarrow}\left([\mathrm{C}]_{\mathrm{eq}} + \delta\right)$$

$$\frac{\mathrm{d}\delta}{\mathrm{d}t} = \underbrace{\left(k_{\rightarrow}[\mathrm{A}]_{\mathrm{eq}}[\mathrm{B}]_{\mathrm{eq}} - k_{\leftarrow}[\mathrm{C}]_{\mathrm{eq}}\right)}_{\alpha=0} - \underbrace{\left[\left(k_{\rightarrow}\left([\mathrm{A}]_{\mathrm{eq}} + [\mathrm{B}]_{\mathrm{eq}}\right) + k_{\leftarrow}\right)\right]}_{\beta}\delta + \underbrace{k_{\rightarrow}}_{\chi}\delta^2$$

E5.

(a)

$$\frac{\mathrm{d}X}{\mathrm{d}t} = k_1 A - k_2 BX + k_3 X^2 Y - k_4 X = Z_1(X,Y) \quad \frac{\mathrm{d}Y}{\mathrm{d}t} = k_2 BX - k_3 X^2 Y = Z_2(X,Y)$$

$$\begin{cases} 0 = k_1 A - k_2 B X_s + k_3 X_s^2 Y_s - k_4 X_s \\ 0 = k_2 B X_s - k_3 X_s^2 Y_s \end{cases} \qquad \begin{cases} X_s = \frac{k_1 A}{k_4} \\ Y_s = \frac{k_2 k_4}{k_1 k_3}\frac{B}{A} \end{cases}$$

(b)

$$\begin{pmatrix} \left(\frac{\partial Z_1}{\partial X}\right)_s & \left(\frac{\partial Z_1}{\partial Y}\right)_s \\ \left(\frac{\partial Z_2}{\partial X}\right)_s & \left(\frac{\partial Z_2}{\partial Y}\right)_s \end{pmatrix} = \begin{pmatrix} -k_2 B - k_4 + 2k_3 X_s Y_s & k_3 X_s^2 \\ k_2 B - 2k_3 X_s Y_s & -k_3 X_s^2 \end{pmatrix}$$

$$= \begin{pmatrix} k_2 B - k_4 & k_3 X_s^2 \\ -k_2 B & -k_3 X_s^2 \end{pmatrix} = \begin{pmatrix} B - 1 & 1 \\ -B & -1 \end{pmatrix}$$

(c) *Mathematica* code and results

```
mm={{B-1,1},{-B,-1}};
Eigenvalues[mm]/.B→2
```

```
{-i, i}
```

E6.

(a)

$$\frac{\mathrm{d}X}{\mathrm{d}t} = k_1AX - k_2XY = Z_1(X,Y) \quad \frac{\mathrm{d}Y}{\mathrm{d}t} = k_2XY - k_3Y = Z_2(X,Y)$$

$$\begin{cases} 0 = k_1AX_s - k_2X_sY_s \\ 0 = k_2X_sY_s - k_3Y_s \end{cases} \quad \begin{cases} X_s = \frac{k_3}{k_2} \\ Y_s = \frac{k_1A}{k_2} \end{cases}$$

(b)

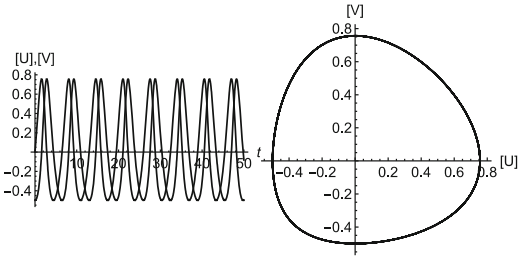
$$\begin{pmatrix} \left(\frac{\partial Z_1}{\partial X}\right)_s & \left(\frac{\partial Z_1}{\partial Y}\right)_s \\ \left(\frac{\partial Z_2}{\partial X}\right)_s & \left(\frac{\partial Z_2}{\partial Y}\right)_s \end{pmatrix} = \begin{pmatrix} k_1A - k_2Y_s & -k_2X_s \\ k_2Y_s & k_2X_s - k_3 \end{pmatrix} = \begin{pmatrix} 0 & -k_3 \\ k_1A & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

(c) *Mathematica* code and results

```
mm={{0,-1},{1,0}};
Eigenvalues[mm]
{i, -i}
```

(d) *Mathematica* code and results

```
k1=1.0; k2=1.0; k3=1.0; A=1.0; u0=0.0; v0=-0.5;
a=k3/k2; b=k1*A/k2; mt=50; ms=10000;
sol=NDSolve[{U'[t]==k1*A*(U[t]+a)-k2*(V[t]+b)*(U[t]+a),
    V'[t]==k2*(V[t]+b)*(U[t]+a)-k3*(V[t]+a),U[0]==u0,V[0]==v0},
    {U,V},{t,0,mt},MaxSteps→ms];
graph=Row[{Plot[Evaluate[{U[t],V[t]}/.sol],{t,0,mt}, AxesLabel → {"t","[U], [V]"},
    AxesStyle→Directive[FontFamily→"Arial",FontSize→14],PlotStyle→{{Thickness[0.006],
    Black},{Thickness[0.006],Black}},PlotRange→All,ImageSize→{250,250}],
    ParametricPlot[Evaluate[{U[t],V[t]}/.sol],{t,0,mt},AxesLabel→{"[U]","[V]"},AxesStyle→
    Directive[FontFamily→"Arial",FontSize→14],PlotStyle→{Thickness[0.006],Black},
    PlotRange→All,ImageSize→{250,250}]}]
```



Chapter 3—The Schrödinger Equation

E1.

$$A\boldsymbol{x} = \lambda \boldsymbol{x} \quad (A - \lambda I)\boldsymbol{x} = 0 \quad \det(A - \lambda I) = |A - \lambda I| = 0$$

where I is the identity matrix (all its elements are equal to zero except those along the main diagonal, which equal one), λ represents the eigenvalues, and $\boldsymbol{x}$ the corresponding eigenvectors.

$$A = \begin{pmatrix} -1 & 1 \\ 4 & 2 \end{pmatrix} \quad \begin{vmatrix} -1-\lambda & 1 \\ 4 & 2-\lambda \end{vmatrix} = 0 \quad \lambda^2 - \lambda - 6 = 0 \quad \lambda_1 = 3 \quad \lambda_2 = -2$$

$$\lambda_1 = 3 \quad \begin{pmatrix} -1 & 1 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 \\ 3x_2 \end{pmatrix} \quad x_2 = 4x_1 \quad \text{eigenvector}\begin{pmatrix} a \\ 4a \end{pmatrix}$$

$$\lambda_2 = -2 \quad \begin{pmatrix} -1 & 1 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix} \quad x_2 = -x_1 \quad \text{eigenvector}\begin{pmatrix} b \\ -b \end{pmatrix}$$

where a and b are constants.

E2. *Mathematica* Code and Results

```
{eigenvals,eigenvecs}=Eigensystem[{{1,0,-1},{1,2,1},{2,2,3}}];
eigenvals
eigenvecs
{3, 2, 1}

{{-1, 1, 2}, {-2, 1, 2}, {-1, 1, 0}}
```

E3.

$$\frac{\partial}{\partial x}(f+g) = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x} \qquad \frac{\partial}{\partial x}(cf) = c\frac{\partial f}{\partial x}$$

E4.

$$\left[\frac{\partial}{\partial x}, x\right] f = \frac{\partial}{\partial x}(xf) - x\frac{\partial f}{\partial x} = f + x\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial x} = f \quad \therefore \quad \left[\frac{\partial}{\partial x}, x\right] = 1$$

E5.

$$\hat{p}_x = \frac{\hbar}{i}\frac{\partial}{\partial x} \quad \frac{p_x^2}{2m} = \frac{1}{2m}\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^2 = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$$

E6.

$$u(x - vt) = u(s) \quad s = x - vt \quad \frac{\partial s}{\partial x} = 1 \quad \frac{\partial s}{\partial t} = -v$$

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial s}\frac{\partial s}{\partial x} = \frac{\partial u}{\partial s} \quad \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial s^2}$$

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial s}\frac{\partial s}{\partial t} = -v\frac{\partial u}{\partial s} \quad \frac{\partial^2 u}{\partial t^2} = -v\frac{\partial\left(-v\frac{\partial u}{\partial s}\right)}{\partial s} = v^2\frac{\partial^2 u}{\partial s^2} \quad \frac{\partial^2 u}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2}$$

If the argument x of the function $u(x)$ is changed by $x - a$, where a is a positive number, then the function $u(x - a)$ is displaced to the right of $u(x)$ by a. If we let $a = vt$, where v is a positive number and t is time, then the function displacement is continuous and increases with time.

E7.

$$u(s) = A\cos(kx - \omega t) \quad s = kx - \omega t \quad \frac{\partial s}{\partial x} = k \quad \frac{\partial s}{\partial t} = -\omega$$

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial s}k = -kA\sin(kx - \omega t) \quad \frac{\partial^2 u}{\partial x^2} = k^2\frac{\partial^2 u}{\partial s^2} = -k^2A\cos(kx - \omega t)$$

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial s}(-\omega) = \omega A\sin(kx - \omega t) \quad \frac{\partial^2 u}{\partial t^2} = \omega^2\frac{\partial^2 u}{\partial s^2} = -\omega^2A\cos(kx - \omega t)$$

$$\frac{1}{k^2}\frac{\partial^2 u}{\partial x^2} = \frac{1}{\omega^2}\frac{\partial^2 u}{\partial t^2} \quad \frac{1}{k^2}k^2\frac{\partial^2 u}{\partial s^2} = \frac{1}{\omega^2}\omega^2\frac{\partial^2 u}{\partial s^2}$$

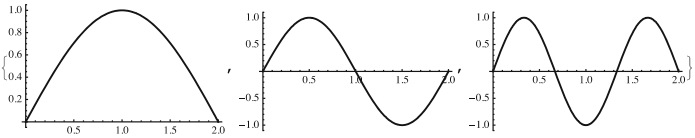
E8. *Mathematica* Code and Results

```
(* d = length of one-dimensional box *)
1/Sqrt[Integrate[Sin[n Pi x/d]^2,{x,0,d}]]
```

$$\frac{2}{\sqrt{d\left(2 - \frac{\text{Sin}[2\,n\,\pi]}{n\,\pi}\right)}}$$

E9. *Mathematica* Code and Results

```
c=2.; (* c = length of one-dimensional box *)
Table[Plot[Sqrt[2/c] Sin[n Pi x/c],{x,0,c},PlotStyle→Black],{n,1,3}]
```



E10. *Mathematica* Code and Results

```
DSolve[y`` [x]+y` [x]-6y[x]==0,y[x],x]
```

```
{{y[x] → e^(-3 x) C[1] + e^(2 x) C[2]}}
```

E11.

$$\int \psi^* \widehat{A}\psi \mathrm{d}\tau = \int \psi \left(\widehat{A}\psi\right)^* \mathrm{d}\tau \quad \psi = f + cg$$

$$\int (f+cg)^* \widehat{A}(f+cg)\mathrm{d}\tau = \int (f+cg)\left[\widehat{A}(f+cg)\right]^* \mathrm{d}\tau$$

$$\int f^* \widehat{A} f \mathrm{d}\tau + c^* c \int g^* \widehat{A} g \mathrm{d}\tau + c^* \int g^* \widehat{A} f \mathrm{d}\tau + c \int f^* \widehat{A} g \mathrm{d}\tau$$

$$= \int f \left(\widehat{A} f\right)^* \mathrm{d}\tau + c^* c \int g \left(\widehat{A} g\right)^* \mathrm{d}\tau + c^* \int f \left(\widehat{A} g\right)^* \mathrm{d}\tau + c \int g \left(\widehat{A} f\right)^* \mathrm{d}\tau$$

Since $\hat{A}$ is Hermitian, the first terms on the left- and right-hand sides of this equation are equal, and so are the second terms. Therefore,

$$c^* \left[\int g^* \widehat{A} f \mathrm{d}\tau - \int f \left(\widehat{A} g\right)^* \mathrm{d}\tau\right] = c \left[\int g \left(\widehat{A} f\right)^* \mathrm{d}\tau - \int f^* \widehat{A} g \mathrm{d}\tau\right]$$

Since c is an arbitrary complex number, the factors in square brackets should be zero. Therefore,

$$\int g^* \widehat{A} f \mathrm{d}\tau = \int f \left(\widehat{A} g\right)^* \mathrm{d}\tau \quad \int f^* \widehat{A} g \mathrm{d}\tau = \int g \left(\widehat{A} f\right)^* \mathrm{d}\tau$$

E12.

$$RM = r_A m_A + r_B m_B \qquad RM = r_A m_A + r_B m_B$$
$$r m_A = -r_A m_A + r_B m_A \qquad -r m_B = r_A m_B - r_B m_B$$
$$r_B = R + r\frac{m_A}{M} \qquad r_A = R - r\frac{m_B}{M}$$
$$\frac{1}{2}m_A\left(\frac{dr_A}{dt}\right)^2 + \frac{1}{2}m_B\left(\frac{dr_B}{dt}\right)^2 = \frac{1}{2}M\left(\frac{dR}{dt}\right)^2 + \frac{1}{2}\mu\left(\frac{dr}{dt}\right)^2$$

E13.

$$\frac{1}{\mu} = \frac{1}{m_A} + \frac{1}{m_B} + \frac{1}{m_C} + \frac{1}{m_D}$$

E14.

$$x = A\sin(\omega t + b) \quad \frac{\mathrm{d}x}{\mathrm{d}t} = \omega A\cos(\omega t + b) \quad \frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\omega^2 A\sin(\omega t + b)$$
$$\omega = \sqrt{\frac{k}{m}} \quad k = m\omega^2 \quad m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -m\omega^2 A\sin(\omega t + b) = -kx$$

E15. *Mathematica* Code and Results

```
Table[HermiteH[n,x],{n,0,5}]
```

```
{1, 2 x, -2 + 4 x^2, -12 x + 8 x^3, 12 - 48 x^2 + 16 x^4, 120 x - 160 x^3 + 32 x^5}
```

E16. *Mathematica* Code and Results

```
aa=Table[1/Sqrt[Integrate[(Exp[-x^2/2]*HermiteH[n,x])^2,
 {x,-Infinity,Infinity}]],{n,0,5}]
bb=Table[1/(Pi^(1/4)Sqrt[2^n n!]),{n,0,5}]
bb-aa
```

$$\left\{\frac{1}{\pi^{1/4}}, \frac{1}{\sqrt{2}\ \pi^{1/4}}, \frac{1}{2\sqrt{2}\ \pi^{1/4}}, \frac{1}{4\sqrt{3}\ \pi^{1/4}}, \frac{1}{8\sqrt{6}\ \pi^{1/4}}, \frac{1}{16\sqrt{15}\ \pi^{1/4}}\right\}$$

$$\left\{\frac{1}{\pi^{1/4}}, \frac{1}{\sqrt{2}\ \pi^{1/4}}, \frac{1}{2\sqrt{2}\ \pi^{1/4}}, \frac{1}{4\sqrt{3}\ \pi^{1/4}}, \frac{1}{8\sqrt{6}\ \pi^{1/4}}, \frac{1}{16\sqrt{15}\ \pi^{1/4}}\right\}$$

$$\{0, 0, 0, 0, 0, 0\}$$

E17.

$$H_{n+1}(z) = 2zH_n(z) - 2nH_{n-1}(z) \quad n = 3 \quad H_4(z) = 2zH_3(z) - 2 \times 3H_2(z)$$
$$H_2(z) = 4z^2 - 2 \quad H_3(z) = 8z^3 - 12z \quad H_4(z) = 16z^4 - 48z^2 + 12$$
$$16z^4 - 48z^2 + 12 = 2z\left(8z^3 - 12z\right) - 6\left(4z^2 - 2\right) = 16z^4 - 48z^2 + 12$$

E18.

$$U(R) = D\left[1 - \mathrm{e}^{-\beta(R-R_e)}\right]^2 \qquad U(R_e) = 0$$
$$\frac{\mathrm{d}U}{\mathrm{d}R} = 2D\beta\left[\mathrm{e}^{-\beta(R-R_e)} - \mathrm{e}^{-2\beta(R-R_e)}\right] \qquad \left(\frac{\mathrm{d}U}{\mathrm{d}R}\right)_{R=R_e} = 0$$
$$\frac{\mathrm{d}^2U}{\mathrm{d}R^2} = 2D\beta^2\left[-\mathrm{e}^{-\beta(R-R_e)} + 2\mathrm{e}^{-2\beta(R-R_e)}\right] \qquad \left(\frac{\mathrm{d}^2U}{\mathrm{d}R^2}\right)_{R=R_e} = 2D\beta^2 = k \quad \beta = \sqrt{\frac{k}{2D}}$$
$$\frac{\mathrm{d}^3U}{\mathrm{d}R^3} = 2D\beta^3\left[\mathrm{e}^{-\beta(R-R_e)} - 4\mathrm{e}^{-2\beta(R-R_e)}\right] \qquad \left(\frac{\mathrm{d}^3U}{\mathrm{d}R^3}\right)_{R=R_e} = -6D\beta^3 = -3k\sqrt{\frac{k}{2D}}$$
$$\frac{\mathrm{d}^4U}{\mathrm{d}R^4} = 2D\beta^4\left[-\mathrm{e}^{-\beta(R-R_e)} + 8\mathrm{e}^{-2\beta(R-R_e)}\right] \qquad \left(\frac{\mathrm{d}^4U}{\mathrm{d}R^4}\right)_{R=R_e} = 14D\beta^4 = \frac{7k^2}{2D}$$

E19.

$$L = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix}$$

$$\boldsymbol{L} = L_x\boldsymbol{i} + L_y\boldsymbol{j} + L_z\boldsymbol{k} = \begin{vmatrix} y & z \\ p_y & p_z \end{vmatrix}\boldsymbol{i} + \begin{vmatrix} z & x \\ p_z & p_x \end{vmatrix}\boldsymbol{j} + \begin{vmatrix} x & y \\ p_x & p_y \end{vmatrix}\boldsymbol{k}$$

$$L_x = y\,p_z - z\,p_y \quad L_y = z\,p_x - x\,p_z \quad L_z = x\,p_y - y\,p_x$$

E20.

Let f be an arbitrary wave function.

$$\begin{aligned} \widehat{L}_x\widehat{L}_y f = \left(y\hat{p}_z - z\hat{p}_y\right)\left(z\hat{p}_x - x\hat{p}_z\right)f &= \left[y\hat{p}_z(z\hat{p}_x) - yx\hat{p}_z^2 - z^2\hat{p}_y\hat{p}_x + zx\hat{p}_y\hat{p}_z\right]f \\ &= \left(\frac{\hbar}{i}y\hat{p}_x + yz\hat{p}_z\hat{p}_x - yx\hat{p}_z^2 - z^2\hat{p}_y\hat{p}_x + zx\hat{p}_y\hat{p}_z\right)f \\ \widehat{L}_y\widehat{L}_x f = \left(z\hat{p}_x - x\hat{p}_z\right)\left(y\hat{p}_z - z\hat{p}_y\right)f &= \left[zy\hat{p}_x\hat{p}_z - z^2\hat{p}_x\hat{p}_y - xy\hat{p}_z^2 + x\hat{p}_z\left(z\hat{p}_y\right)\right]f \\ &= \left(zy\hat{p}_x\hat{p}_z - z^2\hat{p}_x\hat{p}_y - xy\hat{p}_z^2 + \frac{\hbar}{i}x\hat{p}_y + xz\hat{p}_z\hat{p}_y\right)f \end{aligned}$$

Mixed partial second derivatives are equal, for example, $\partial^2 f/\partial x\partial y = \partial^2 f/\partial y\partial x$. Thus,

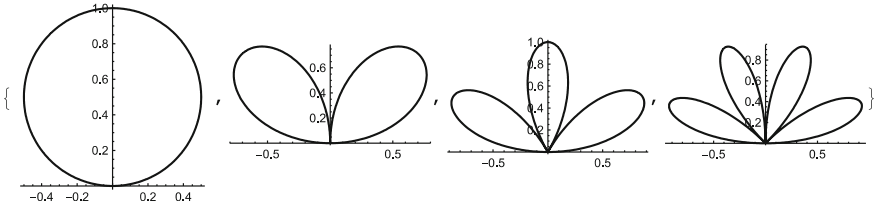
$$\left(\widehat{L}_x\widehat{L}_y - \widehat{L}_y\widehat{L}_x\right)f = \frac{\hbar}{i}\left(y\hat{p}_x - x\hat{p}_y\right)f = -\frac{\hbar}{i}\widehat{L}_z f = i\hbar\widehat{L}_z f \quad \left[\widehat{L}_x, \widehat{L}_y\right] = i\hbar\widehat{L}_z$$

E21.

$$\begin{aligned}
\left[\widehat{L}^2,\,\widehat{L}_x\right] &= \left[\widehat{L}_x^2+\widehat{L}_y^2+\widehat{L}_z^2,\,\widehat{L}_x\right] = \left[\widehat{L}_y^2,\,\widehat{L}_x\right]+\left[\widehat{L}_z^2,\,\widehat{L}_x\right] \\
&= \widehat{L}_y\widehat{L}_y\widehat{L}_x-\widehat{L}_x\widehat{L}_y\widehat{L}_y+\widehat{L}_z\widehat{L}_z\widehat{L}_x-\widehat{L}_x\widehat{L}_z\widehat{L}_z \\
&= \left(\widehat{L}_y\widehat{L}_y\widehat{L}_x-\underline{\widehat{L}_y\widehat{L}_x\widehat{L}_y}\right)+\left(\underline{\widehat{L}_y\widehat{L}_x\widehat{L}_y}-\widehat{L}_x\widehat{L}_y\widehat{L}_y\right) \\
&\quad+\left(\widehat{L}_z\widehat{L}_z\widehat{L}_x-\underline{\underline{\widehat{L}_z\widehat{L}_x\widehat{L}_z}}\right)+\left(\underline{\underline{\widehat{L}_z\widehat{L}_x\widehat{L}_z}}-\widehat{L}_x\widehat{L}_z\widehat{L}_z\right) \\
&= \widehat{L}_y\left[\widehat{L}_y,\,\widehat{L}_x\right]+\left[\widehat{L}_y,\,\widehat{L}_x\right]\widehat{L}_y+\widehat{L}_z\left[\widehat{L}_z,\,\widehat{L}_x\right]+\left[\widehat{L}_z,\,\widehat{L}_x\right]\widehat{L}_z \\
&= -i\hbar\widehat{L}_y\widehat{L}_z-i\hbar\widehat{L}_z\widehat{L}_y+i\hbar\widehat{L}_z\widehat{L}_y+i\hbar\widehat{L}_y\widehat{L}_z \\
&= 0
\end{aligned}$$

E22. *Mathematica* Code and Results

```
Table[PolarPlot[Abs[Sin[nθ]],{θ,0,Pi},PlotStyle→Black,
  ImageSize→{150,150}],{n,1,4}]
```

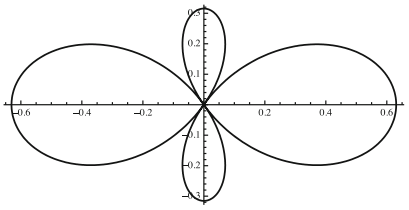


When the parameter n increases, more points are sampled, and the function changes more rapidly.

E23. *Mathematica* Code and Results

```
p20=SphericalHarmonicY[2,0,θ,0]
PolarPlot[Abs[p20],{θ,0,2 Pi},PlotStyle→Black]
```

$$\frac{1}{4}\sqrt{\frac{5}{\pi}}\,\left(-1+3\,\mathrm{Cos}[\theta]^2\right)$$



E24.

$$\widehat{L}^2 = -\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right) \quad \widehat{L}_z = -i\hbar\frac{\partial}{\partial\phi}$$

$$\widehat{H} = -\frac{\hbar^2}{2\mu}\nabla^2 + V(r) \quad \nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{1}{r^2\hbar^2}\widehat{L}^2$$

$$\left[\widehat{H}, \widehat{L}^2\right] = \left[-\frac{\hbar^2}{2\mu}\nabla^2 + V(r), \widehat{L}^2\right] = \left[-\frac{\hbar^2}{2\mu}\nabla^2, \widehat{L}^2\right] + \left[V(r), \widehat{L}^2\right] = 0 + 0 = 0$$

$$\left[\widehat{H}, \widehat{L}_z\right] = \left[-\frac{\hbar^2}{2\mu}\nabla^2 + V(r), \widehat{L}_z\right] = \left[-\frac{\hbar^2}{2\mu}\nabla^2, \widehat{L}_z\right] + \left[V(r), \widehat{L}_z\right] = 0 + 0 = 0$$

E25.

The square of the radial factor is a probability density, thus having units of inverse of volume, that is, a_0^{-3}, where $a_0 = 0.52917721092 \times 10^{-10}$ m is the atomic unit of length. Therefore, all radial factors come with the factor $a_0^{-3/2}$. In SI units, we have $R_{2s} = \frac{a_0^{-3/2}}{2\sqrt{2}}\left(2 - \frac{r}{a_0}\right)\mathrm{e}^{-r/(2a_0)}$. In atomic units, $a_0 = 1$ and $R_{2s} = \frac{1}{2\sqrt{2}}(2 - r)\mathrm{e}^{-r/2}$.

E26. *Mathematica* Code and Results

```
Simplify[m1*r1^2+m2*r2^2-(m1*m2/(m1+m2)*(r1+r2)^2,m1*r1-m2*r2==0]
0
```

E27.

$$W[\phi] = \frac{\int \phi^* \widehat{H}\,\phi\,\mathrm{d}\tau}{\int \phi^*\,\phi\,\mathrm{d}\tau} = \frac{\Lambda}{\Delta} \quad \phi = \sum_{j=1}^{n} c_j f_j$$

$$\frac{\partial W}{\partial c_j} = 0 \quad \frac{\partial\left(\frac{\Lambda}{\Delta}\right)}{\partial c_j} = \frac{\Delta\left(\frac{\partial\Lambda}{\partial c_j}\right) - \Lambda\left(\frac{\partial\Delta}{\partial c_j}\right)}{\Delta^2} = \frac{\frac{\partial\Lambda}{\partial c_j} - W\frac{\partial\Delta}{\partial c_j}}{\Delta} = 0 \quad \frac{\partial\Lambda}{\partial c_j} - W\frac{\partial\Delta}{\partial c_j} = 0$$

E28. *Mathematica* Code and Results

```
Solve[x+y-z==2&&2x-y+3z==5&&3x+2y-2z==5,{x,y,z}]
```

```
{{x → 1, y → 3, z → 2}}
```

E29. *Mathematica* Code and Results

```
m={{a,b,c},{d,e,f},{g,h,i}};
n={{r,s,t},{u,v,x},{y,z,w}};
p=m.n;
Row[{m//MatrixForm,".", n//MatrixForm,"=",p//MatrixForm},""]
```

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \cdot \begin{pmatrix} r & s & t \\ u & v & x \\ y & z & w \end{pmatrix} = \begin{pmatrix} ar+bu+cy & as+bv+cz & at+cw+bx \\ dr+eu+fy & ds+ev+fz & dt+fw+ex \\ gr+hu+iy & gs+hv+iz & gt+iw+hx \end{pmatrix}$$

E30.

$$\begin{vmatrix} (H_{11}-WS_{11}) & (H_{12}-WS_{12}) \\ (H_{12}-WS_{12}) & (H_{22}-WS_{22}) \end{vmatrix} = 0 \quad \frac{(H_{11}-WS_{11})}{(H_{12}-WS_{12})} = \frac{(H_{12}-WS_{12})}{(H_{22}-WS_{22})}$$

$$\begin{cases} (H_{11}-WS_{11})c_1 + (H_{12}-WS_{12})c_2 = 0 \\ (H_{12}-WS_{12})c_1 + (H_{22}-WS_{22})c_2 = 0 \end{cases}$$

$$\begin{cases} \dfrac{(H_{11}-WS_{11})}{(H_{12}-WS_{12})}c_1 + c_2 = 0 \\ \dfrac{(H_{12}-WS_{12})}{(H_{22}-WS_{22})}c_1 + c_2 = 0 \end{cases} \quad \begin{cases} \dfrac{c_2}{c_1} = -\dfrac{(H_{11}-WS_{11})}{(H_{12}-WS_{12})} \\ \dfrac{c_2}{c_1} = -\dfrac{(H_{12}-WS_{12})}{(H_{22}-WS_{22})} \end{cases}$$

E31.

$$V_1 = \sum_{j=2}^{n} V_{1j} = \sum_{j=2}^{n} \int \frac{|g_j|^2}{r_{1j}} \mathrm{d}v_j$$

E32.

Because equation (3.154) applies to molecules. A molecule has more than one nucleus, and V_{NN} represents the nucleus–nucleus repulsion energy. This energy term does not exist for atoms, because they have one nucleus.

E33.

When applied to atoms, (3.154) does not include the energy term V_{NN}. Note that the sums in (3.154) are over doubly occupied orbitals. From the definitions of Coulomb and exchange integrals, (3.155) and (3.156), and the fact that electron labels 1 and 2 are dummy variables, it can be inferred that $J_{ij} = J_{ji}$ and $K_{ij} = K_{ji}$. In addition, $K_{ii} = J_{ii}$. Therefore, applying (3.154) to the $1s^2 2s^2$ electron configuration leads to

$$E_{\mathrm{HF}} = 2(\varepsilon_{1s} + \varepsilon_{2s}) \; - J_{1s1s} - J_{2s2s} - 2(2J_{1s2s} - K_{1s2s})$$

E34.

$$K_{ij} = \iint \frac{\psi_i^*(1)\psi_j(1)\psi_j^*(2)\psi_i(2)}{r_{12}} \mathrm{d}\tau_1 \mathrm{d}\tau_2$$

If $\psi_i = \phi_i\alpha$ and $\psi_j = \phi_j\beta$

$$\begin{aligned} K_{ij} &= \iint \frac{\psi_i^*(1)\psi_j(1)\psi_j^*(2)\psi_i(2)}{r_{12}} \mathrm{d}\tau_1 \mathrm{d}\tau_2 \\ &= \iint \frac{\phi_i^*(1)\phi_j(1)\phi_j^*(2)\phi_i(2)}{r_{12}} \mathrm{d}v_1 \mathrm{d}v_2 \int \alpha^*(1)\beta(1)\mathrm{d}\sigma_1 \int \beta^*(2)\alpha(2)\mathrm{d}\sigma_2 = 0 \end{aligned}$$

due to integration over the spin variables.

E35.

$$J_{ij} = \left\langle \phi_i(1)\phi_j(2) \middle| \frac{1}{r_{12}} \middle| \phi_i(1)\phi_j(2) \right\rangle \quad K_{ij} = \left\langle \phi_i(1)\phi_j(2) \middle| \frac{1}{r_{12}} \middle| \phi_j(1)\phi_i(2) \right\rangle$$

E36. *Mathematica* Code and Results

```
f[r_]:=Exp[-ζ r]*r
norm=1./Sqrt[Integrate[f[r]^2 r^2,{r,0,Infinity},Assumptions→ζ>0]]
(* verify norm *)
Integrate[norm^2*f[r]^2 r^2,{r,0,Infinity},Assumptions→ζ>0]
```

$$\frac{1.1547}{\sqrt{\frac{1}{\zeta^5}}}$$

```
1.
```

E37.

The square of the absolute value of a Slater-type orbital is a probability density, thus having units of the inverse of volume, that is, a_0^{-3}. Therefore, all radial factors come with the factor $a_0^{-3/2}$. In SI units,

$$N_{n,\zeta}\left(\frac{r}{a_0}\right)^{n-1} \mathrm{e}^{-\zeta r/a_0} \quad \text{where} \quad N_{n,\zeta} = \frac{(2\zeta)^{n+1/2}}{[(2n)!]^{1/2}} \frac{1}{a_0^{3/2}}$$

In atomic units, $a_0 = 1$, and the radial factor of a Slater-type orbital is given by

$$N_{n,\zeta} r^{n-1} \mathrm{e}^{-\zeta r} \quad \text{where} \quad N_{n,\zeta} = \frac{(2\zeta)^{n+1/2}}{[(2n)!]^{1/2}}$$

E38.

$$\Psi = \phi(1)\phi(2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)]/\sqrt{2}$$

$$\begin{aligned}\rho_0(1) &= \int \Psi^*\Psi \, \mathrm{d}v_2 \mathrm{d}\sigma_1 \mathrm{d}\sigma_2 \\ &= \int \phi^*(1)\phi^*(2)\phi(1)\phi(2)\mathrm{d}v_2 \frac{1}{2}\int [\alpha^*(1)\beta^*(2) \\ &\quad - \alpha^*(2)\beta^*(1)][\alpha(1)\beta(2) - \alpha(2)\beta(1)]\mathrm{d}\sigma_1 \mathrm{d}\sigma_2 \\ &= \phi^*(1)\phi(1)\frac{1}{2}(1 - 0 - 0 + 1) = |\phi|^2\end{aligned}$$

E39.

$$\begin{aligned}\Psi =& \phi(1)\phi(2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)]/\sqrt{2} \\ &\int \Psi^*\left[\widehat{T}(1) + \widehat{T}(2)\right]\Psi \mathrm{d}v_1 \mathrm{d}v_2 \mathrm{d}\sigma_1 \mathrm{d}\sigma_2 = 2\int \phi^*(1)\widehat{T}(1)\phi(1)\mathrm{d}v_1\end{aligned}$$

E40.

$$\hat{H}_0 = -\frac{\nabla^2}{2} - \frac{1}{r} \quad \nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}\hat{L}^2 \quad \hat{L}^2 Y_{\ell m} = \ell(\ell+1)Y_{\ell m} \quad \hat{H}' = -z$$

$$R_{1s} = 2\mathrm{e}^{-r} \quad Y_{0,0} = \frac{1}{2\sqrt{\pi}} \quad 1s = \frac{1}{\sqrt{\pi}}\mathrm{e}^{-r} \quad 2p_z(\mathrm{STO}) = R_{2\zeta}Y_{1,0}$$

$$= \left(\frac{2\zeta^{5/2}}{\sqrt{3}} r\mathrm{e}^{-\zeta r}\right)\left(\frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\theta\right) = \frac{1}{\sqrt{\pi}}\zeta^{5/2} r\mathrm{e}^{-\zeta r}\cos\theta$$

$$z = r\cos\theta \quad \langle 1s|-z|2p_z\rangle = -\frac{\zeta^{5/2}}{\pi}\int_0^\infty \mathrm{e}^{-(\zeta+1)r} r^4 \mathrm{d}r \int_0^\pi \cos^2\theta \sin\theta \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi$$

$$= -\frac{4}{3}\zeta^{5/2}\int_0^\infty \mathrm{e}^{-(\zeta+1)r} r^4 \mathrm{d}r = -\frac{32\sqrt{\zeta^5}}{(1+\zeta)^5}$$

$$\left\langle 1s\middle|-\frac{\nabla^2}{2} - \frac{1}{r}\middle|1s\right\rangle = -0.5 \quad \left\langle R_{2\zeta}\middle|-\frac{\nabla^2}{2} - \frac{1}{r}\middle|R_{2\zeta}\right\rangle$$

$$= -\frac{2\zeta^5}{3}\int_0^\infty \mathrm{e}^{-2\zeta r}(\zeta^2 r^2 - 4\zeta r + 2r)\mathrm{d}r = -\frac{1}{6}(2 - 3\zeta)\zeta^3$$

$$E_0^{(1)} = \langle 1s|-z|1s\rangle = 0 \quad J = \left\langle R_{2\zeta}\middle|\hat{H}^0\middle|R_{2\zeta}\right\rangle - \left\langle 1s\middle|\hat{H}^0\middle|1s\right\rangle$$

$$+ 2\langle 1s|-z|2p_z\rangle \geq E_0^{(2)} \quad \min J(\zeta = 0.844)$$

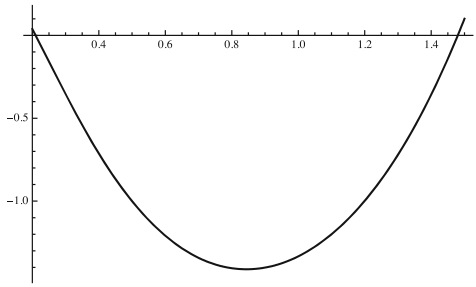
Mathematica code and results:

```
h12=-(4/3) Sqrt[ζ^5] Integrate[x^4 Exp[-(1+ζ) x],{x,0,Infinity},
  Assumptions→ ζ >0]
h22=-(2/3) ζ^5 Integrate[Exp[-2ζ x] (ζ^2 x^2-4ζ x+2x)]],{x,0,Infinity},
  Assumptions→ ζ >0]
h11=-0.5;
ζ/.SetPrecision[FindMinimum[h22-h11+2h12,{ζ,0.6}][[2]],3]
Plot[h22-h11+2h12,{ζ,0,1.5},PlotStyle→Black]
```

$$-\frac{32\sqrt{\zeta^5}}{(1+\zeta)^5}$$

$$-\frac{1}{6}(2-3\zeta)\zeta^3$$

```
0.844
```



E41.

$$s = \omega_{nm} \pm \omega \quad \lim_{s\to 0}\frac{e^{is\tau}-1}{s} = \frac{d(e^{is\tau}-1)/ds}{ds/ds} = i\tau$$

E42.

The lowest and highest wavelengths of the visible region are approximately equal to 400 and 800 nm. The 400 nm wavelength is $2^4 = 16$ times more intensely scattered than the wavelength at 800 nm. This explains the color of the sky.

E43. *Mathematica* Code and Results

```
alfa={{axx,axy,axz},{axy,ayy,ayz},{axz,ayz,azz}};
Print["Trace of the polarizability tensor"]
aa=alfa[[1,1]]+alfa[[2,2]]+alfa[[3,3]]
Print["Square of the polarizability anisotropy"]
g2=((alfa[[1,1]]-alfa[[2,2]])^2+(alfa[[2,2]]-alfa[[3,3]])^2+(alfa[[3,3]]-
alfa[[1,1]])^2+6*(alfa[[1,2]]^2+alfa[[2,3]]^2+alfa[[3,1]]^2))/2
xx={{Cos[x],-Sin[x],0},{Sin[x],Cos[x],0},{0,0,1}};
beta=Simplify[Transpose[xx].alfa.xx];
bb=beta[[1,1]]+beta[[2,2]]+beta[[3,3]];
h2=((beta[[1,1]]-beta[[2,2]])^2+(beta[[2,2]]-beta[[3,3]])^2+(beta[[3,3]]-
beta[[1,1]])^2+6*(beta[[1,2]]^2+beta[[2,3]]^2+beta[[3,1]]^2))/2;
Print["Effect of an orthogonal transformation on the trace"]
Simplify[bb-aa]
Print["Effect of an orthogonal transformation on the anisotropy"]
Simplify[h2-g2]
```

```
Trace of the polarizability tensor

axx + ayy + azz

Square of the polarizability anisotropy

1/2 ((axx - ayy)^2 + 6 (axy^2 + axz^2 + ayz^2) + (ayy - azz)^2 + (-axx + azz)^2)

Effect of an orthogonal transformation on the trace

0

Effect of an orthogonal transformation on the anisotropy

0
```

E44.

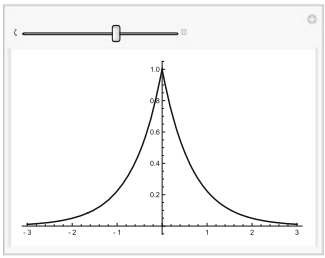
The intensity of the Stokes line is proportional to the population of level a in the vibrational transition $a \rightarrow b$, whereas the intensity of the corresponding anti-Stokes line is proportional to the population of the excited level b in the vibrational transition $b \rightarrow a$. Hence, using the Boltzmann distribution and v^4-dependence of the Stokes and anti-Stokes lines, we can write

$$\rho_b = \rho_a \exp(-hv_{ba}/kT) \quad v_{ba} = v_b - v_a$$
$$\frac{I_{\text{Stokes}}}{I_{\text{anti-Stokes}}} = \frac{(v_0 - v_{ba})^4}{(v_0 + v_{ba})^4}\frac{\rho_a}{\rho_b} = \frac{(v_0 - v_{ba})^4}{(v_0 + v_{ba})^4}\exp(hv_{ba}/kT)$$

It can be concluded that the ratio of the Stokes and anti-Stokes intensities increases rapidly as v_{ba} increases.

E45. *Mathematica* Code and Results

```
Manipulate[Plot[Exp[-c*Abs[x]],{x,-3,3},PlotStyle→Black],{{c,1.5,"ζ"},0.7,2.0,0.1}]
```



E46.

For C_2H_6 at 298 K, $E_{\text{MP2/6-311G*}} - E_{\text{HF/6-311G*}} \approx -177.7$ kcal mol^{-1}, whereas for the chemical reaction $C_2H_4(g) + H_2(g) \rightarrow C_2H_6(g)$ at 298 K, $E_{\text{MP2/6-311G*}} - E_{\text{HF/6-311G*}} \approx 1.7$ kcal mol^{-1}. The MP2 correlation correction for the ethane molecule corresponds to 0.36 % of the Hartree–Fock energy and is about the energy of two C–H bonds. For the chemical reaction, the MP2 correlation correction is less than 1 % of the absolute value of the corresponding value for the molecule. This finding suggests that the correlation energy is mainly due to the electrons of each electron pair. Now, since the number of electron bond pairs is conserved for reactants and products, the correlation correction is appreciably lower than for a molecule. In addition, the correlation correction for a reaction is not necessarily negative, since it is the correlation correction for the products minus the correlation correction for the reactants.

Chapter 4—Molecular Symmetry

E1.

(a) C_3 (b) C_3 (c) C_4 (d) C_5 (e) C_{2v} (f) D_{3h} (*g*) D_{6h} (h) $D_{\infty h}$ (i) T_d

E2. *Mathematica* Code and Results

```
cn={{Cos[θ],-Sin[θ],0},{Sin[θ],Cos[θ],0},{0,0,1}};
r1=cn//MatrixForm;
r2=Transpose[cn]//MatrixForm;
r3=(cn.Transpose[cn])//MatrixForm;
r4=Simplify[cn.Transpose[cn]]//MatrixForm;
Row[{r1,".",r2,"=",r3,"=",r4},""]
```

$$\begin{pmatrix} \cos[\theta] & -\sin[\theta] & 0 \\ \sin[\theta] & \cos[\theta] & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos[\theta] & \sin[\theta] & 0 \\ -\sin[\theta] & \cos[\theta] & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos[\theta]^2+\sin[\theta]^2 & 0 & 0 \\ 0 & \cos[\theta]^2+\sin[\theta]^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

E3. *Mathematica* Code and Results

```
x=Pi/2;
c4={{Cos[x],-Sin[x],0},{Sin[x],Cos[x],0},{0,0,1}};
r={{1,0,0},{0,1,0},{0,0,-1}};
s4=r.c4;
a1=r//MatrixForm;
a2=c4//MatrixForm;
a3=s4//MatrixForm;
Row[{a1,".",a2,"=",a3},""]
```

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

E4. *Mathematica* Code and Results

```
aa={{a11,a12,a13},{a21,a22,a23},{a31,a32,a33}};
bb={{b11,b12,b13},{b21,b22,b23},{b31,b32,b33}};
(Transpose[aa.bb]-Transpose[bb].Transpose[aa])//MatrixForm
```

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

E5. *Mathematica* Code and Results

```
aa={{a11,a12,a13},{a21,a22,a23},{a31,a32,a33}};
bb={{b11,b12,b13},{b21,b22,b23},{b31,b32,b33}};
Simplify[Inverse[aa.bb]-Inverse[bb].Inverse[aa]]//MatrixForm
```

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

E6.

$$\sum_k B_{kk} = \sum_k \sum_{i,j} \left(X^{-1}\right)_{ki} A_{ij} X_{jk} = \sum_k \sum_{i,j} A_{ij} X_{jk} \left(X^{-1}\right)_{ki} = \sum_{i,j} A_{ij} \left(X X^{-1}\right)_{ji}$$
$$= \sum_{i,j} A_{ij} \delta_{ji} = \sum_i A_{ii}$$

(QED)

E7.

(a)

$$A = X^{-1} A X \Rightarrow X A X^{-1} = A$$

(b)

$$B = X^{-1} A X \Rightarrow X B X^{-1} = A$$

(c)

$$A = X^{-1} B X \quad A = Y^{-1} C Y$$
$$Y^{-1} C Y = X^{-1} B X \quad C = Y X^{-1} B X Y^{-1}$$
$$C = \left(XY^{-1}\right)^{-1} B \left(XY^{-1}\right) \quad C = Z^{-1} B Z$$

E8. *Mathematica* Code and Results

```
e={{1,0,0},{0,1,0},{0,0,1}};
c3={{Cos[2Pi/3],-Sin[2Pi/3],0},{Sin[2Pi/3],Cos[2Pi/3],0},{0,0,1}};
c32={{Cos[4Pi/3],-Sin[4Pi/3],0},{Sin[4Pi/3],Cos[4Pi/3],0},{0,0,1}};
sv1={{-1,0,0},{0,1,0},{0,0,1}};
sv2=Transpose[c3].sv1.c3;
sv3=Transpose[c32].sv1.c32;

a1,1="E";a1,2="C3"; a1,3="C3^2";a1,4="σv1";a1,5="σv2"; a1,6="σv3";
a2,1="C3"; a3,1="C3^2";a4,1="σv1";a5,1="σv2"; a6,1="σv3";

If[c3.c3==sv1,a2,2="σv1",If[c3.c3==sv2,a2,2="σv2",If[c3.c3==sv3,a2,2="σv3",
    If[c3.c3==c3,a2,2="C3",If[c3.c3==c32,a2,2="C3^2",If[c3.c3==e,a2,2="E",f]]]]]];
If[c3.c32==sv1,a2,3="σv1",If[c3.c32==sv2,a2,3="σv2",If[c3.c32==sv3,a2,3="σv3",
    If[c3.c32==c3,a2,3="C3",If[c3.c32==c32,a2,3="C3^2",If[c3.c32==e,a2,3="E",f]]]]]];
If[c3.sv1==sv1,a2,4="σv1",If[c3.sv1==sv2,a2,4="σv2",If[c3.sv1==sv3,a2,4="σv3",
    If[c3.sv1==c3,a2,4="C3",If[c3.sv1==c32,a2,4="C3^2",If[c3.sv1==e,a2,4="E",f]]]]]];
If[c3.sv2==sv1,a2,5="σv1",If[c3.sv2==sv2,a2,5="σv2",If[c3.sv2==sv3,a2,5="σv3",
    If[c3.sv2==c3,a2,5="C3",If[c3.sv2==c32,a2,5="C3^2",If[c3.sv2==e,a2,5="E",f]]]]]];
If[c3.sv3==sv1,a2,6="σv1",If[c3.sv3==sv2,a2,6="σv2",If[c3.sv3==sv3,a2,6="σv3",
    If[c3.sv3==c3,a2,6="C3",If[c3.sv3==c32,a2,6="C3^2",If[c3.sv3==e,a2,6="E",f]]]]]];
```

```
If[c32.c3==sv1,a3,2="σv1",If[c32.c3==sv2,a3,2="σv2",If[c32.c3==sv3,a3,2="σv3",
    If[c32.c3==c3,a3,2="C3",If[c32.c3==c32,a3,2="C3^2",If[c32.c3==e,a3,2="E",f]]]]]];
If[c32.c32==sv1,a3,3="σv1",If[c32.c32==sv2,a3,3="σv2",If[c32.c32==sv3,a3,3="σv3",
    If[c32.c32==c3,a3,3="C3",If[c32.c32==c32,a3,3="C3^2",If[c32.c32==e,a3,3="E",f]]]]]];
If[c32.sv1==sv1,a3,4="σv1",If[c32.sv1==sv2,a3,4="σv2",If[c32.sv1==sv3,a3,4="σv3",
    If[c32.sv1==c3,a3,4="C3",If[c32.sv1==c32,a3,4="C3^2",If[c32.sv1==e,a3,4="E",f]]]]]];
If[c32.sv2==sv1,a3,5="σv1",If[c32.sv2==sv2,a3,5="σv2",If[c32.sv2==sv3,a3,5="σv3",
    If[c32.sv2==c3,a3,5="C3",If[c32.sv2==c32,a3,5="C3^2",If[c32.sv2==e,a3,5="E",f]]]]]];
If[c32.sv3==sv1,a3,6="σv1",If[c32.sv3==sv2,a3,6="σv2",If[c32.sv3==sv3,a3,6="σv3",
    If[c32.sv3==c3,a3,6="C3",If[c32.sv3==c32,a3,6="C3^2",If[c32.sv3==e,a3,6="E",f]]]]]];

If[sv1.c3==sv1,a4,2="σv1",If[sv1.c3==sv2,a4,2="σv2",If[sv1.c3==sv3,a4,2="σv3",
    If[sv1.c3==c3,a4,2="C3",If[sv1.c3==c32,a4,2="C3^2",If[sv1.c3==e,a4,2="E",f]]]]]];
If[sv1.c32==sv1,a4,3="σv1",If[sv1.c32==sv2,a4,3="σv2",If[sv1.c32==sv3,a4,3="σv3",
    If[sv1.c32==c3,a4,3="C3",If[sv1.c32==c32,a4,3="C3^2",If[sv1.c32==e,a4,3="E",f]]]]]];
If[sv1.sv1==sv1,a4,4="σv1",If[sv1.sv1==sv2,a4,4="σv2",If[sv1.sv1==sv3,a4,4="σv3",
    If[sv1.sv1==c3,a4,4="C3",If[sv1.sv1==c32,a4,4="C3^2",If[sv1.sv1==e,a4,4="E",f]]]]]];
If[sv1.sv2==sv1,a4,5="σv1",If[sv1.sv2==sv2,a4,5="σv2",If[sv1.sv2==sv3,a4,5="σv3",
    If[sv1.sv2==c3,a4,5="C3",If[sv1.sv2==c32,a4,5="C3^2",If[sv1.sv2==e,a4,5="E",f]]]]]];
If[sv1.sv3==sv1,a4,6="σv1",If[sv1.sv3==sv2,a4,6="σv2",If[sv1.sv3==sv3,a4,6="σv3",
    If[sv1.sv3==c3,a4,6="C3",If[sv1.sv3==c32,a4,6="C3^2",If[sv1.sv3==e,a4,6="E",f]]]]]];

If[sv2.c3==sv1,a5,2="σv1",If[sv2.c3==sv2,a5,2="σv2",If[sv2.c3==sv3,a5,2="σv3",
    If[sv2.c3==c3,a5,2="C3",If[sv2.c3==c32,a5,2="C3^2",If[sv2.c3==e,a5,2="E",f]]]]]];
If[sv2.c32==sv1,a5,3="σv1",If[sv2.c32==sv2,a5,3="σv2",If[sv2.c32==sv3,a5,3="σv3",
    If[sv2.c32==c3,a5,3="C3",If[sv2.c32==c32,a5,3="C3^2",If[sv2.c32==e,a5,3="E",f]]]]]];
If[sv2.sv1==sv1,a5,4="σv1",If[sv2.sv1==sv2,a5,4="σv2",If[sv2.sv1==sv3,a5,4="σv3",
    If[sv2.sv1==c3,a5,4="C3",If[sv2.sv1==c32,a5,4="C3^2",If[sv2.sv1==e,a5,4="E",f]]]]]];
If[sv2.sv2==sv1,a5,5="σv1",If[sv2.sv2==sv2,a5,5="σv2",If[sv2.sv2==sv3,a5,5="σv3",
    If[sv2.sv2==c3,a5,5="C3",If[sv2.sv2==c32,a5,5="C3^2",If[sv2.sv2==e,a5,5="E",f]]]]]];
If[sv2.sv3==sv1,a5,6="σv1",If[sv2.sv3==sv2,a5,6="σv2",If[sv2.sv3==sv3,a5,6="σv3",
    If[sv2.sv3==c3,a5,6="C3",If[sv2.sv3==c32,a5,6="C3^2",If[sv2.sv3==e,a5,6="E",f]]]]]];

If[sv3.c3==sv1,a6,2="σv1",If[sv3.c3==sv2,a6,2="σv2",If[sv3.c3==sv3,a6,2="σv3",
    If[sv3.c3==c3,a6,2="C3",If[sv3.c3==c32,a6,2="C3^2",If[sv3.c3==e,a6,2="E",f]]]]]];
If[sv3.c32==sv1,a6,3="σv1",If[sv3.c32==sv2,a6,3="σv2",If[sv3.c32==sv3,a6,3="σv3",
    If[sv3.c32==c3,a6,3="C3",If[sv3.c32==c32,a6,3="C3^2",If[sv3.c32==e,a6,3="E",f]]]]]];
If[sv3.sv1==sv1,a6,4="σv1",If[sv3.sv1==sv2,a6,4="σv2",If[sv3.sv1==sv3,a6,4="σv3",
    If[sv3.sv1==c3,a6,4="C3",If[sv3.sv1==c32,a6,4="C3^2",If[sv3.sv1==e,a6,4="E",f]]]]]];
If[sv3.sv2==sv1,a6,5="σv1",If[sv3.sv2==sv2,a6,5="σv2",If[sv3.sv2==sv3,a6,5="σv3",
    If[sv3.sv2==c3,a6,5="C3",If[sv3.sv2==c32,a6,5="C3^2",If[sv3.sv2==e,a6,5="E",f]]]]]];
If[sv3.sv3==sv1,a6,6="σv1",If[sv3.sv3==sv2,a6,6="σv2",If[sv3.sv3==sv3,a6,6="σv3",
    If[sv3.sv3==c3,a6,6="C3",If[sv3.sv3==c32,a6,6="C3^2",If[sv3.sv3==e,a6,6="E",f]]]]]];


m=Grid[Array[a## &,{6,6}],Frame→All,ItemSize→{2,2}]
```

E	C_3	C_3^2	σ_{v1}	σ_{v2}	σ_{v3}
C_3	C_3^2	E	σ_{v2}	σ_{v3}	σ_{v1}
C_3^2	E	C_3	σ_{v3}	σ_{v1}	σ_{v2}
σ_{v1}	σ_{v3}	σ_{v2}	E	C_3^2	C_3
σ_{v2}	σ_{v1}	σ_{v3}	C_3	E	C_3^2
σ_{v3}	σ_{v2}	σ_{v1}	C_3^2	C_3	E

This point group is not abelian (not commutative).

E9.

Infrared active vibrations transform like x, y, and z, and for a point group with a center of inversion, these functions are of plus–minus- or u-type. In turn, Raman active vibrations transform like the quadratic functions x^2, y^2, z^2, xy, xz, and yz, and for a point group with a center of inversion, these functions are of plus–plus- or g-type.

E10.

C_{3v} E $2C_3$ $3\sigma_v$

Γ 15 0 3

$\Gamma = 4A_1 + A_2 + 5E$

$\Gamma_{rot} = A_2 + E$ $\Gamma_{trans} = A_1 + E$

$\Gamma_{vib} = 3A_1 + 3E$

C_{2v} E C_2 σ_{xz} σ_{yz}

Γ 15 −1 3 3

$\Gamma = 5A_1 + 2A_2 + 4B_1 + 4B_2$

$\Gamma_{rot} = A_2 + B_1 + B_2$ $\Gamma_{trans} = A_1 + B_1 + B_2$

$\Gamma_{vib} = 4A_1 + A_2 + 2B_1 + 2B_2$

C_s E σ_h

Γ 15 3

$\Gamma = 9A' + 6A''$

$\Gamma_{rot} = A' + 2A''$ $\Gamma_{trans} = 2A' + A''$

$\Gamma_{vib} = 6A' + 3A''$

C_1 E

Γ 15

$\Gamma = 15A$

$\Gamma_{rot} = 3A$ $\Gamma_{trans} = 3A$

$\Gamma_{vib} = 9A$

Chapter 5—Molecular Structure

E1.

$$\phi(P) = \sum_n \frac{Z_n}{r_{nP}} - \int \frac{\rho_0(\boldsymbol{r})}{r_{eP}} \mathrm{d}v$$

The first term of this expression is the electrostatic potential, which results from the discrete distribution of nuclei of the molecule. The second term results from the continuous distribution of electronic charge, where $\rho_0(\boldsymbol{r})\mathrm{d}v$ is the electronic charge on the volume element $\mathrm{d}v$ centered at P.

E2.

$$\Psi_0(1,2) = \frac{1}{\sqrt{2}}|\phi_1\alpha(1)\phi_1\beta(2)| = \frac{1}{\sqrt{2}}\phi_1(1)\phi_1(2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)]$$

$$\phi_1 = \mathrm{MO} \quad \phi_1 = c_{11}\chi_1 + c_{21}\chi_2 \quad \chi_1, \chi_2 = \mathrm{AOs} \quad \chi_1 = 1s_1 \quad \chi_2 = 1s_2$$

$$\rho_0(\boldsymbol{r}) = 2\int \Psi_0^*(1,2)\Psi_0(1,2)\mathrm{d}v_2\mathrm{d}\sigma_1\mathrm{d}\sigma_2 \quad \rho_0(\boldsymbol{r}) = 2\phi_1^*\phi_1 = 2c_{11}^2\chi_1^2 + 2c_{21}^2\chi_2^2 + 4c_{11}c_{21}\chi_1\chi_2$$

$$\int \rho_0(\boldsymbol{r})\mathrm{d}v = 2c_{11}^2 + 2c_{21}^2 + 4c_{11}c_{21}S_{12} = 2 \quad S_{12} = \int \chi_1\chi_2\mathrm{d}v$$

$$c_{11}^2 + c_{21}^2 + 2c_{11}c_{21}S_{12} = 1 \quad c_{11}^2 = c_{21}^2 \quad |c_{11}| = |c_{21}| = \frac{1}{\sqrt{2(1+S_{12})}}$$

E3. *Mathematica* Code and Results

```
Print["CH4 HF/STO-3G: coeffs of 5 double occ. MOs (= columns) in terms of AOs"]
c={{0.99193,-0.22143,0,0,0},{0.03836,0.62759,0,0,0},{0,0,0,0,0.57197},
  {0,0,0.57197,0,0},{0,0,0,0.57197,0},{-0.00701,0.18057,0.30082,0.30082,0.30082},
  {-0.00701,0.18057,-0.30082,0.30082,-0.30082},
  {-0.00701,0.18057,0.30082,-0.30082,-0.30082},
  {-0.00701,0.18057,-0.30082,-0.30082,0.30082}}
c//MatrixForm
Print["Density matrix"]
p=2 c.Transpose[c];
Chop[SetPrecision[p//MatrixForm,4],10^-5]
```

```
CH4 HF/STO-3G: coefficients of 5 double occ. MOs(=columns) in terms of AOs
```

$$\begin{pmatrix} 0.99193 & -0.22143 & 0 & 0 & 0 \\ 0.03836 & 0.62759 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.57197 \\ 0 & 0 & 0.57197 & 0 & 0 \\ 0 & 0 & 0 & 0.57197 & 0 \\ -0.00701 & 0.18057 & 0.30082 & 0.30082 & 0.30082 \\ -0.00701 & 0.18057 & -0.30082 & 0.30082 & -0.30082 \\ -0.00701 & 0.18057 & 0.30082 & -0.30082 & -0.30082 \\ -0.00701 & 0.18057 & -0.30082 & -0.30082 & 0.30082 \end{pmatrix}$$

```
Density matrix
```

$$\begin{pmatrix} 2.066 & -0.2018 & 0 & 0 & 0 & -0.09387 & -0.09387 & -0.09387 & -0.09387 \\ -0.2018 & 0.7907 & 0 & 0 & 0 & 0.2261 & 0.2261 & 0.2261 & 0.2261 \\ 0 & 0 & 0.6543 & 0 & 0 & 0.3441 & -0.3441 & -0.3441 & 0.3441 \\ 0 & 0 & 0 & 0.6543 & 0 & 0.3441 & -0.3441 & 0.3441 & -0.3441 \\ 0 & 0 & 0 & 0 & 0.6543 & 0.3441 & 0.3441 & -0.3441 & -0.3441 \\ -0.09387 & 0.2261 & 0.3441 & 0.3441 & 0.3441 & 0.6083 & -0.1157 & -0.1157 & -0.1157 \\ -0.09387 & 0.2261 & -0.3441 & -0.3441 & 0.3441 & -0.1157 & 0.6083 & -0.1157 & -0.1157 \\ -0.09387 & 0.2261 & -0.3441 & 0.3441 & -0.3441 & -0.1157 & -0.1157 & 0.6083 & -0.1157 \\ -0.09387 & 0.2261 & 0.3441 & -0.3441 & -0.3441 & -0.1157 & -0.1157 & -0.1157 & 0.6083 \end{pmatrix}$$

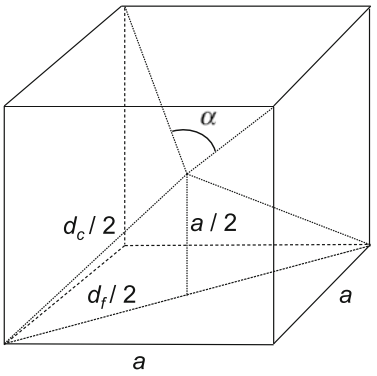
E4.

Full Mulliken population matrix element between $1s_2$ and $1s_3$: −0.01993.
Density matrix element between $1s_2$ and $1s_3$: −0.11568.
Overlap matrix element between $1s_2$ and $1s_3$: −0.01993/−0.11568 = 0.172286.

E5.

Population condensed to C atom: 2.06590 + 0.79067 + 0.65429 × 3 + (−0.05013 × 2) = 4.71918.
Population condensed to the overlap region between C and H_1 atoms: −0.00594 + 0.11187 + 0.09331 × 3 = 0.38586.

E6.



$$d_f^2 = \mathrm{a}^2 + \mathrm{a}^2 \quad d_f = \sqrt{2}\,\mathrm{a}$$

$$\left(\frac{d_c}{2}\right)^2 = \left(\frac{d_f}{2}\right)^2 + \left(\frac{\mathrm{a}}{2}\right)^2 \quad d_c = \sqrt{3}\,\mathrm{a}$$

$$\frac{d_f}{2} = \frac{d_c}{2}\sin\frac{\alpha}{2} \quad \sin\frac{\alpha}{2} = \sqrt{\frac{2}{3}} \quad \alpha = 2\arcsin\sqrt{\frac{2}{3}} \quad \alpha = 109.4712^\circ$$

E7.

$$h_A = \tfrac{1}{2}(s + p_x + p_y + p_z) \qquad h_B = \tfrac{1}{2}(s - p_x + p_y - p_z)$$
$$h_C = \tfrac{1}{2}(s + p_x - p_y - p_z) \qquad h_D = \tfrac{1}{2}(s - p_x - p_y + p_z)$$
$$S_{AA} = \tfrac{1}{4}(1 + 1 + 1 + 1) = 1 \qquad S_{AA} = S_{BB} = S_{CC} = S_{DD} = 1$$
$$S_{AB} = \tfrac{1}{4}(1 - 1 + 1 - 1) = 0 \qquad S_{AB} = S_{AC} = S_{AD} = S_{BC} = S_{BD} = S_{CD} = 0$$

E8. *Mathematica* Code and Results

```
c3={{Cos[2Pi/3],-Sin[2Pi/3]},{Sin[2Pi/3],Cos[2Pi/3]}};
c32=c3.c3;
pA=px;
pB=(c3.{{px},{py}})[[1,1]];
pC=(c32.{{px},{py}})[[1,1]];
hA=1/Sqrt[3] (s+Sqrt[2] pA);
hB=1/Sqrt[3] (s+Sqrt[2] pB);
hC=1/Sqrt[3] (s+Sqrt[2] pC);
{hA,hB,hC}
```

$$\left\{\frac{\sqrt{2}\,\text{px}+\text{s}}{\sqrt{3}},\ \frac{\sqrt{2}\left(-\frac{\text{px}}{2}-\frac{\sqrt{3}\,\text{py}}{2}\right)+\text{s}}{\sqrt{3}},\ \frac{\sqrt{2}\left(-\frac{\text{px}}{2}+\frac{\sqrt{3}\,\text{py}}{2}\right)+\text{s}}{\sqrt{3}}\right\}$$

$$h_A=\frac{1}{\sqrt{3}}(s+\sqrt{2}p_x)\quad h_B=\frac{1}{\sqrt{3}}\left[s+\sqrt{2}\left(-\frac{p_x}{2}-\frac{\sqrt{3}p_y}{2}\right)\right]$$

$$h_C=\frac{1}{\sqrt{3}}\left[s+\sqrt{2}\left(-\frac{p_x}{2}+\frac{\sqrt{3}p_y}{2}\right)\right]$$

$$S_{AB}=\int h_A h_B \mathrm{d}v=\frac{1}{3}\left[1+2\left(-\frac{1}{2}-0\right)\right]=0\quad S_{AB}=S_{AC}=S_{BC}=0$$

$$S_{AA}=S_{BB}=S_{CC}=1$$

E9.

$$h_A=\frac{1}{\sqrt{2}}(s+p_x)\quad h_B=\frac{1}{\sqrt{2}}(s-p_x)$$

$$S_{AB}=\int h_A h_B \mathrm{d}v=\frac{1}{2}(1-1)=0\quad S_{AA}=S_{BB}=1$$

E10.

$$h_A = \frac{1}{\sqrt{1+\lambda_A}}\left(s + \sqrt{\lambda_A}p_A\right) \quad h_B = \frac{1}{\sqrt{1+\lambda_B}}\left(s + \sqrt{\lambda_B}p_B\right)$$

$$S_{h_A h_B} = \frac{1}{\sqrt{(1+\lambda_A)(1+\lambda_B)}}\left(S_{ss} + \sqrt{\lambda_A\lambda_B}S_{p_A p_B} + \sqrt{\lambda_A}S_{p_A s} + \sqrt{\lambda_B}S_{sp_B}\right)$$

$$= \frac{1}{\sqrt{(1+\lambda_A)(1+\lambda_B)}}\left(1 + \sqrt{\lambda_A\lambda_B}S_{p_A p_B} + \sqrt{\lambda_A}\times 0 + \sqrt{\lambda_B}\times 0\right) = 0$$

$$1 + \sqrt{\lambda_A\lambda_B}S_{p_A p_B} = 0 \quad S_{p_A p_B} = -\frac{1}{\sqrt{\lambda_A\lambda_B}}$$

E11.

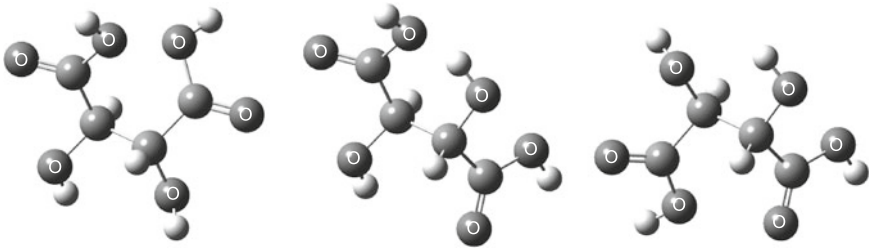
1 = antiperiplanar conformer ($g_1 = 1$) 2 = synclinal conformer ($g_2 = 2$)

$$f_2 \propto g_2 e^{-E_2/(RT)} = 2e^{-E_2/(RT)} \quad f_1 \propto e^{-E_1/(RT)} \quad \frac{f_2}{f_1} = 2e^{-(E_2-E_1)/(RT)}$$

$$\ln\left(\frac{f_2}{2f_1}\right) = -\frac{E_2 - E_1}{RT} \quad \Delta E_{21} = E_2 - E_1 = RT\ln\left(\frac{2f_1}{f_2}\right)$$

$$= 2.5\ln\left(\frac{2\times 0.80}{0.20}\right) = 5.2 \text{ kJ mol}^{-1}$$

E12.

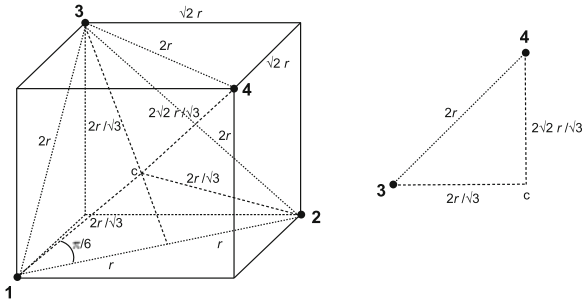


R,R *R,S* *S,S*

The naturally occurring form is the *R*,*R* stereoisomer. The *R*,*S* stereoisomer is optically inactive.

Chapter 6—Crystals

E1.



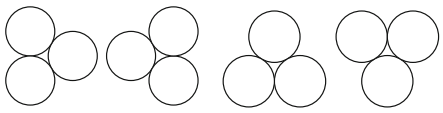
Comparing with Fig. 6.6, all distances now have a scaling factor $2r$.

E2.

Considering Fig. 6.6, the distance from point c (the intersection of the threefold axis with the plane of the disks) to the disk centers is equal to $d = 1/\sqrt{3}$. Hence, $x = d\cos\theta$ and $y = d\sin\theta$. Note that when $r = 1/2$, $d = 1$.

Mathematica code and results:

```
aa=Graphics[Table[{EdgeForm[{Thickness[0.004],Black}],White,
   Disk[{Cos[2πk/3]/Sqrt[3],Sin[2πk/3]/Sqrt[3]},1/2]},{k,0,2}],ImageSize→250];
bb=Graphics[Table[{EdgeForm[{Thickness[0.004],Black}],White,
   Disk[{-Cos[2πk/3]/Sqrt[3],-Sin[2πk/3]/Sqrt[3]},1/2]},{k,0,2}],ImageSize→250];
cc=Graphics[Table[{EdgeForm[{Thickness[0.004],Black}],White,
   Disk[{Sin[2πk/3]/Sqrt[3],Cos[2πk/3]/Sqrt[3]},1/2]},{k,0,2}],ImageSize→250];
dd=Graphics[Table[{EdgeForm[{Thickness[0.004],Black}],White,
   Disk[{-Sin[2πk/3]/Sqrt[3],-Cos[2πk/3]/Sqrt[3]},1/2]},{k,0,2}],ImageSize→250];
Row[{" ",aa," ",bb," ",cc," ",dd}]
c1=Table[{Cos[2πk/3]/Sqrt[3],Sin[2πk/3]/Sqrt[3]},{k,0,2}];
c2=Table[{-Cos[2πk/3]/Sqrt[3],-Sin[2πk/3]/Sqrt[3]},{k,0,2}];
c3=Table[{Sin[2πk/3]/Sqrt[3],Cos[2πk/3]/Sqrt[3]},{k,0,2}];
c4=Table[{-Sin[2πk/3]/Sqrt[3],-Cos[2πk/3]/Sqrt[3]},{k,0,2}];
Row[{c1," ",c2," ",c3," ",c4}]
```



$\{\{\frac{1}{\sqrt{3}}, 0\}, \{-\frac{1}{2\sqrt{3}}, \frac{1}{2}\}, \{-\frac{1}{2\sqrt{3}}, -\frac{1}{2}\}\}$ $\{\{-\frac{1}{\sqrt{3}}, 0\}, \{\frac{1}{2\sqrt{3}}, -\frac{1}{2}\}, \{\frac{1}{2\sqrt{3}}, \frac{1}{2}\}\}$

$\{\{0, \frac{1}{\sqrt{3}}\}, \{\frac{1}{2}, -\frac{1}{2\sqrt{3}}\}, \{-\frac{1}{2}, -\frac{1}{2\sqrt{3}}\}\}$ $\{\{0, -\frac{1}{\sqrt{3}}\}, \{-\frac{1}{2}, \frac{1}{2\sqrt{3}}\}, \{\frac{1}{2}, \frac{1}{2\sqrt{3}}\}\}$

Inspection of the above (x,y) coordinates shows that disks are drawn in the counterclockwise direction in the first two graphs and in the clockwise direction in the last two graphs.

E3.

Consider a hexagonal prism as repeating unit (see Fig. 6.8). Each sphere in a prism vertex is shared by six repeating units, thus being considered with a factor 1/6 in the packing density. The centers of the bottom and top of the hexagonal prism are occupied by one sphere that is shared by two repeating units, thus having the factor 1/2. Note that there are three spheres inside the hexagonal prism counting 1 each of them. The volume of the spheres in one repeating unit is given by

$$\left[6 \times \frac{1}{6} + 6 \times \frac{1}{6} + 2 \times \frac{1}{2} + 3\right] \frac{4}{3}\pi r^3 = 8\pi r^3$$

In turn, the volume of the hexagonal prism can be obtained by adding the volume of three parallelepipeds whose height is given by $2 \times 2(\sqrt{2}/\sqrt{3})r$ (see answer to **E1**). The edge of the base parallelogram of each parallelepiped is given by $2r$, which multiplies by $r\sqrt{3}$. Therefore, the volume of the hexagonal prism is given by

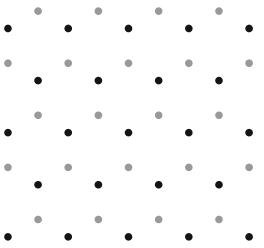
$$3 \times 2r \times r\sqrt{3} \times 2 \times \frac{2\sqrt{2}}{\sqrt{3}} r = 3 \times 8 \times \sqrt{2} r^3$$

The packing density of the hcp structure of identical spheres is obtained by dividing the volume of the spheres contained within one hexagonal prism as repeating unit of the hcp structure by the total volume of the hexagonal prism,

$$\frac{8\pi r^3}{3 \times 8 \times \sqrt{2}\, r^3} = \frac{\pi}{3\sqrt{2}} \approx 0.74$$

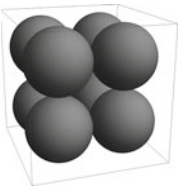
E4. *Mathematica* Code and Results

```
v1={1,0}; v2={1/2,Sqrt[3]/2};t=(v1+v2)/3;
listA= Flatten[Table[n1 v1+n2 v2,{n1,-3,3},{n2,-3,3}],1];
listB= Flatten[Table[n1 v1+n2 v2+t,{n1,-3,3},{n2,-3,3}],1];
latticeA=Graphics[{{Black,PointSize[0.03],Point[listA]}},PlotRange→{{-2.1,2.1},{-2.1,2.1}},
  ImageSize→300,Axes→False];
latticeB=Graphics[{{Gray,PointSize[0.03],Point[listB]}}];
hex2Dmerge=Show[latticeA,latticeB]
```



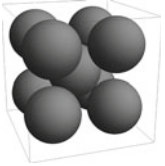
E5. *Mathematica* Code and Results

```
r=0.5; a=2r;
v1=a{1,0,0}; v2=a{0,1,0}; v3=a{0,0,1};
list=Flatten[Table[n1 v1+n2 v2+n3 v3,{n1,-2,2},{n2,-2,2},{n3,-2,2}],2];
lattice=Pick[list,Cases[list,{x_,y_,z_}→(x≤1&&x≥0)&&(y≤1&&y≥0)&&(z≤1&&z≥0)]];
sc=Graphics3D[{Opacity[1],Gray,Sphere[lattice,r]},Lighting→"Neutral",
  ImageSize→250,ViewPoint→{3,-1,1}]
```



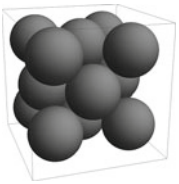
E6. *Mathematica* Code and Results

```
r=0.5; a=4r/Sqrt[3];
v1=a{1,0,0}; v2=a{0,1,0}; v3=a/2{1,1,1};
list=Flatten[Table[n1 v1+n2 v2+n3 v3,{n1,-2,2},{n2,-2,2},{n3,-2,2}],2];
lattice=Pick[list,Cases[list,{x_,y_,z_}→(x<1.16&&x≥0)&&(y<1.16&&y≥0)&&(z<1.16&&z≥0)]];
bcc=Graphics3D[{Opacity[1],Gray,Sphere[lattice,r]},
  Lighting→"Neutral",ImageSize→250,ViewPoint→{3,-1,1}]
```



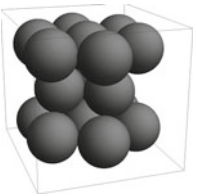
E7. *Mathematica* Code and Results

```
r=0.5; a=2r Sqrt[2];
v1=(a/2){0,1,1}; v2=(a/2){1,0,1}; v3=(a/2){1,1,0};
list=Flatten[Table[n1 v1+n2 v2+n3 v3,{n1,-2,2},{n2,-2,2},{n3,-2,2}],2];
lattice=Pick[list,Cases[list,{x_,y_,z_}→(x<1.42&&x≥0)&&(y<1.42&&y≥0)&&(z<1.42&&z≥0)]];
ccp=Graphics3D[{Opacity[1],Gray,Sphere[lattice,r]},Lighting→"Neutral",
  ImageSize→250,ViewPoint→{3,-1,1}]
```



E8. *Mathematica* Code and Results

```
r=0.5; a=2r; c=4r Sqrt[2/3];
v1=a{1,0,0}; v2=a{1/2,Sqrt[3]/2,0}; v3=c{0,0,1}; t=v1/3+v2/3+v3/2;
list1=Flatten[Table[n1 v1+n2 v2+n3 v3,{n1,-2,2},{n2,-2,2},{n3,-2,2}],2];
list2=Flatten[Table[n1 v1+n2 v2+n3 v3+t,{n1,-2,2},{n2,-2,2},{n3,-2,2}],2];
lattice1=Pick[list1,Cases[list1,{x_,y_,z_}→(x≤2.5&&x≥0.5)&&(y<1.74&&y≥0)&&(z<1.64&&z≥0)]];
lattice2=Pick[list2,Cases[list2,{x_,y_,z_}→(x≤2&&x≥1)&&(y<1.16&&y>0.28)&&
    (z<0.82&&z>0.81)]];
hcp=Graphics3D[{{Opacity[1],Gray,Sphere[lattice1,r]},{Opacity[1],Gray,
    Sphere[lattice2,r]}},Lighting→"Neutral",ImageSize→250,ViewPoint→{3,-1,1}]
```



E9. *Mathematica* Code and Results

```
(* Pi/(N+1) = 2a *)
Simplify[TrigExpand[Cos[2a*(k+1)]-Cos[2a*k]+2Sin[a]*Sin[a*(2k+1)]]]
```

```
0
```

E10.

$$\phi(\varepsilon) = \frac{1}{1 + \exp[(\varepsilon - \mu)/(k_B T)]} \Rightarrow \varepsilon = \mu + k_B T \ln\left(\frac{1}{\phi} - 1\right)$$

$$\mu = 3.24 \text{ eV} \quad T = 7000 \text{ K} \Rightarrow \varepsilon = 3.24 + k_B 7000 \ln\left(\frac{1}{0.25} - 1\right) = 3.90 \text{ eV}.$$

Chapter 7—Water

E1.

The definition of the dipole moment vector is given by $\boldsymbol{\mu} = \sum_k q_k \boldsymbol{r}_k$.

For a pure electric dipole, that is, two electric charges $+q$ and $-q$ separated by a distance δ, the above expression is consistent with the dipole moment vector being directed from the negative to the positive charge. In fact, if we take the origin of the dipole moment at the positive charge, the positional vector $\boldsymbol{r}_+$ is zero ($q\boldsymbol{r}_+ = 0$), and the positional vector that points to the negative charge $\boldsymbol{r}_-$ is multiplied by the negative charge $-q$, thus giving a dipole moment vector $\boldsymbol{\mu} = q\boldsymbol{r}_+ + (-q)\boldsymbol{r}_- = -q\boldsymbol{r}_-$ that points to the positive charge.

E2.

$$\sum q_i \boldsymbol{r}'_i = \sum q_i (\boldsymbol{r}_i + \boldsymbol{R}) = \sum q_i \boldsymbol{r}_i + \boldsymbol{R} \underbrace{\sum q_i}_{0} = \sum q_i \boldsymbol{r}_i$$

where $\boldsymbol{R}$ is the vector that changes the origin.

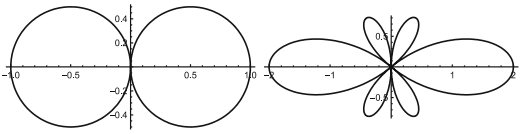
E3.

$$\mu(\mathrm{H_2O}) = 2 \times 0.340 \times 0.9611 \times \cos(104.53/2 \times \pi/180) \times 4.8032 = 1.92 \text{ D}.$$

E4. *Mathematica* Code and Results

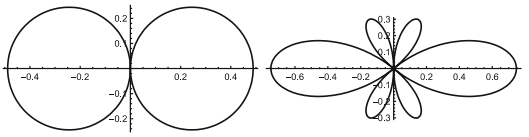
```
aa=PolarPlot[Abs[Cos[θ]],{θ,0,2Pi},PlotStyle→Black,ImageSize→{250,250}];
bb=PolarPlot[Abs[5 Cos[θ]^3-3 Cos[θ]],{θ,0,2Pi},PlotStyle→Black,ImageSize→{250,250}];
cc=PolarPlot[Abs[SphericalHarmonicY[1,0,θ,0]],{θ,0,2Pi},
  PlotStyle→Black,ImageSize→{250,250}];
dd=PolarPlot[Abs[SphericalHarmonicY[3,0,θ,0]],{θ,0,2Pi},
  PlotStyle→Black,ImageSize→{250,250}];
Print["Polar Plots of cos[θ] and 5cos^3[θ]-3cos[θ]"]
Row[{aa,bb},"  "]
Print["Spherical Harmonics in the plane Y_1^0[θ,0] and Y_3^0[θ,0]"]
Row[{cc,dd},"  "]
```

PolarPlots of $\cos[\theta]$ and $5\cos^3[\theta]-3\cos[\theta]$



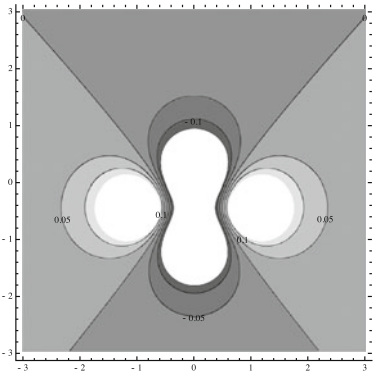
Spherical Harmonics in the plane $Y_1^0[\theta,0]$ and $Y_3^0[\theta,0]$

$$\left\{\frac{1}{2}\sqrt{\frac{3}{\pi}}\,\text{Cos}[\theta],\ \frac{1}{4}\sqrt{\frac{7}{\pi}}\,\left(-3\,\text{Cos}[\theta]+5\,\text{Cos}[\theta]^3\right)\right\}$$



E5. *Mathematica* Code and Results

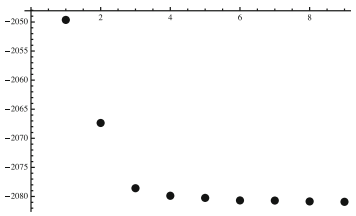
```
elp[q_,p_,r_]:=Sum[q[[i]]/Norm[r-p[[i]]],{i,Length[q]}]
c=Table[Lighter[Gray,i],{i,0,1,1/6}];
quadrupole=ContourPlot[elp[{-0.225,-0.225,0.225,0.225},{{0.,0.1176},{0.,-1.05922},
    {-1.,-0.48},{1.,-0.48}},{x,y}],{x,-3,3},{y,-3,3},ContourShading→c,ContourLabels→True]
```



E6. *Mathematica* Code and Results

```
(* data for B3LYP calculations on H2O taken from CCCBDB *)
basis={"STO-3G","3-21G","6-31G","6-311G*","6-311G**","6-311+G(3df,2p)","6-311+G(3df,2pd)",
  "cc-pVQZ","aug-cc-pVQZ"};
energy=27.211385*{-75.322774,-75.973965,-76.386117,-76.433933,-76.447448,-76.463325,
  -76.464155,-76.469646,-76.472240}; (* eV *)
data = Partition[Riffle[basis, energy], 2]
ListPlot[energy,PlotStyle→{Black,PointSize[0.025]}]
```

```
{{STO-3G, -2049.64}, {3-21G, -2067.36}, {6-31G, -2078.57},
 {6-311G*, -2079.87}, {6-311G**, -2080.24}, {6-311+G(3df,2p), -2080.67},
 {6-311+G(3df,2pd), -2080.7}, {cc-pVQZ, -2080.84}, {aug-cc-pVQZ, -2080.92}}
```



This graph shows that for the same molecule (H_2O) and method of calculation (B3LYP), the increase of the basis set lowers the ground state energy, as stated by the variational principle (see Sect.3.7).

E7.

$$\omega = \sqrt{\frac{k}{\mu}} \quad \mu_{OD} = \frac{16 \times 2}{18}\,\mu_{OH} = \frac{16 \times 1}{17} \quad \frac{\omega_{OH}}{\omega_{OD}} \approx \sqrt{\frac{\mu_{OD}}{\mu_{OH}}} \approx \sqrt{\frac{32 \times 17}{18 \times 16}} \approx \sqrt{2}$$

E8.

H_2O belongs to the C_{2v} point group. The antisymmetric stretching transforms as B_2; the symmetric stretching and the bending transform as A_1.

DHO belongs to the C_s point group. The normal modes are totally symmetric (transform as A′) because all displacement vectors are in the molecular plane.

E9.

$$\begin{array}{ll}(35,35,35,35) & 3^4 = 81\\ (35,35,35,37)\times 4 & 3^3\times 4 = 108\\ (35,35,37,37)\times 6 & 3^2\times 6 = 54\\ (35,37,37,37)\times 4 & 3\times 4 = 12\\ (37,37,37,37) & 1 = 1\end{array}$$

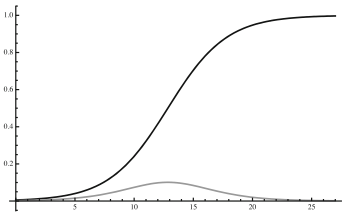
Relative intensities of the distinct isotopic lines: 81, 108, 54, 12, 1.

$$\begin{array}{ll}(35,35,35,35) & \frac{1}{\mu} = \frac{1}{35}+\frac{1}{35}+\frac{1}{35}+\frac{1}{35} \quad \mu = 8.75\\ (35,35,35,37) & \frac{1}{\mu'} = \frac{1}{35}+\frac{1}{35}+\frac{1}{35}+\frac{1}{37} \quad \mu' \approx 8.87\end{array}$$

$$\frac{\omega'}{\omega} = \sqrt{\frac{\mu}{\mu'}} \quad |\omega' - \omega| = \left|459\sqrt{\frac{8.75}{8.87}} - 459\right| \approx 3\ \mathrm{cm}^{-1}$$

E10. *Mathematica* Code and Results

```
R=0.0083145; a=178;
f[x_,T_]:=1/(1+a E^(-x/(RT)));
g[x_,T_]:=D[f[x,T],x];
Plot[{f[x,298.15], g[x,298.15]},{x,0,27},PlotStyle→{{Black,Thickness[0.005]},
  {Gray,Thickness[0.005]},PlotRange→All,ImageSize→{250,250}]
sol=SetPrecision[FindMaximum[g[x,298.15],{x,13.}],3][[2,1]];
x/.sol
```




```
12.8
```

E11.

For a diluted solution,

$$x_s = \frac{n_s}{n_s + n_S} \approx \frac{n_s}{n_S} = \frac{n_s / 1\ \mathrm{mol\ dm}^{-3}}{n_S / 1\ \mathrm{mol\ dm}^{-3}} = \frac{c_s}{c_S} = c_s V_S$$

where c_s is the solute molarity, c_S is the solvent molarity (for water, $c_S = 55.55$ mol dm^{-3}). Note that the inverse of c_S is the molar volume of the solvent in dm^3 (for water, $V_m = 0.018$ dm^3 mol^{-1}).

E12.

$$I = (1/2)\sum_k z_k^2\left(m_k/m^0\right) = (1/2)\left[2^2 \times 0.001 + (-1)^2 \times 0.002\right] = 0.003$$

$$\log\gamma_{\text{salt}} = -A|z_c z_a|I^{1/2} = -0.509 \times 2 \times 0.003^{1/2} = -0.05576 \quad \gamma_{\text{salt}} = 0.880$$

E13.

$T = 298.15$ K $\quad p_{O_2} = p_{H_2} = p^0 = 1$ atm $\quad a_{H_2O} \approx 1$

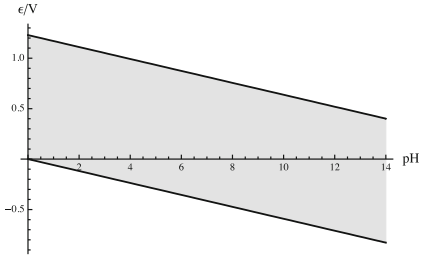
$$O_2 + 4H^+ + 4e^- \rightarrow 2H_2O \quad \varepsilon_{O_2/H_2O} = \varepsilon^0_{O_2/H_2O} - \frac{RT}{4F}\ln\frac{a^2_{H_2O}}{(p_{O_2}/p^0)\,a^4_{H^+}} = 1.229 - 0.05916\ \text{pH}$$

$$2H^+ + 2e^- \rightarrow H_2 \quad \varepsilon_{H^+/H_2} = -\frac{RT}{2F}\ln\frac{p_{H_2}/p^0}{a^2_{H^+}} = -0.05916\ \text{pH}$$

An oxidizing agent that converts H_2O into O_2 and a reducing agent that converts H_2O to H_2 are not thermodynamically stable in aqueous solution. Redox couples that are thermodynamically stable in water lie between the above sloping straight lines $\varepsilon = \varepsilon(\text{pH})$ (region marked gray).

Mathematica code and results:

```
f[x_]:=1.229-0.05916 x
g[x_]:=-0.05916 x
Plot[{f[x],g[x]},{x,0,14},AxesOrigin→{0,0},AxesLabel→{Style["pH",Medium,Black],
  Style["ε/V",Medium,Black]},PlotStyle→{Black,Black},Filling→{1→{2}},
  FillingStyle→LightGray]
```



E14.

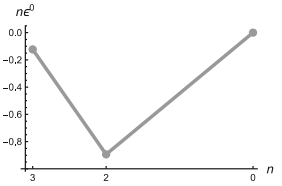
$$2\varepsilon^0_{42} = \varepsilon^0_{43} + \varepsilon^0_{32} \quad \varepsilon^0_{43} = 2\varepsilon^0_{42} - \varepsilon^0_{32} = 2 \times 1.224 - 1.5415 = 0.907$$

E15.

$$3\varepsilon^0_{74} = \varepsilon^0_{76} + 2\varepsilon^0_{64} \quad 2\varepsilon^0_{64} = 3\varepsilon^0_{74} - \varepsilon^0_{76} = 3 \times 1.679 - 0.558 = 2.240$$

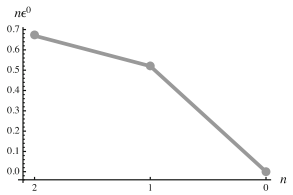
E16. *Mathematica* Code and Results

```
(* values for acidic aqueous solution taken from CRC Handbook of Chemistry and Physics *)
(* Fe3+ → 0.771 V → Fe2+ → -0.447 V → Fe *)
iron=ListPlot[{{-3,-0.123},{-2,-0.894},{0,0.0}],AxesLabel→{Style["n",Medium,Black],
  Style["nε0",Medium,Black]},AxesOrigin→{-3.1,-1.0},PlotStyle→{Gray,PointSize[0.035]},
  Ticks→{{{0,0},{-2,2},{-3,3}},True},ImageSize→{250,250}];
brokenlineiron=ListPlot[{{-3,-0.123},{-2,-0.894},{0,0.0}],Joined→True,
  PlotRange→{{-3.1,0.1},{-0.95,0.05}},PlotStyle→{Gray,Thickness[0.015]}];
Show[iron,brokenlineiron]
```



E17. *Mathematica* Code and Results

```
(* values for acidic aqueous solution taken from CRC Handbook of Chemistry and Physics *)
(* Cu2+ → 0.153 V → Cu+ → 0.521 V → Cu *)
copper=ListPlot[{{-2,0.674},{-1,0.521},{0,0.0}],AxesLabel→{Style["n",Medium,Black],
  Style["nε0",Medium,Black]},AxesOrigin→{-2.1,-0.04},PlotStyle→{Gray,PointSize[0.035]},
  Ticks→{{{0,0},{-1,1},{-2,2}},True},ImageSize→{250,250}];
brokenlinecopper=ListPlot[{{-2,0.674},{-1,0.521},{0,0.0}],Joined→True,
  PlotRange→{{-2.1,0.15},{-0.02,0.70}},PlotStyle→{Gray,Thickness[0.015]}];
Show[copper,brokenlinecopper]
```



E18.

Temperature is assumed to be 298.15 K, and subscripts c and a stand for cathode and anode. Galvani cells have positive $\Delta\varepsilon$.

$$\varepsilon_c = \frac{RT}{F}\ln a_{\mathrm{Ag}^+(c)} = 0.05916\ \log a_{\mathrm{Ag}^+(c)} \quad \varepsilon_a = \frac{RT}{F}\ln a_{\mathrm{Ag}^+(a)} = 0.05916\ \log a_{\mathrm{Ag}^+(a)}$$

$$\varepsilon_c > \varepsilon_a \quad \Delta\varepsilon = \varepsilon_c - \varepsilon_a = 0.05916\ \log\frac{a_{\mathrm{Ag}^+(c)}}{a_{\mathrm{Ag}^+(a)}} > 0 \quad \therefore \quad a_{\mathrm{Ag}^+(c)} > a_{\mathrm{Ag}^+(a)}$$

Index

J.J.C. Teixeira-Dias, *Molecular Physical Chemistry*,
DOI 10.1007/978-3-319-41093-7