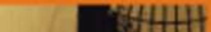
mathematics for chemistry and physics



Mathematics for CHEMISTRY AND PHYSICS

GEORGE TURRELL

$$y(x,0) = \sum_{n=1}^{\infty} B_n \sin$$

Mathematics for Chemistry and Physics

This Page Intentionally Left Blank

Mathematics for Chemistry and Physics

GEORGE TURRELL
University of Science and Technology, Lille, France

ACADEMIC PRESS

An Elsevier Science Imprint

San Diego San Francisco New York Boston
London Sydney Tokyo

This book is printed on acid-free paper.

Academic Press
An Elsevier Science Imprint
Harcourt Place, 32 Jamestown Road, London NW1 7BY, UK
http://www.academicpress.com

Academic Press
An Elsevier Science Imprint
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
http://www.academicpress.com

ISBN 0-12-705051-5

Library of Congress Catalog Number: 2001 091916

A catalogue record for this book is available from the British Library

Typeset by Laserwords Pvt. Ltd., Chennai, India
Printed and bound in Great Britain by MPG Books Ltd, Bodmin, Cornwall

02 03 04 05 06 07 MP 9 8 7 6 5 4 3 2 1

Contents

9 Molecular Mechanics 215

10 Probability and Statistics 245

Appendices

Preface

This book has been written in an attempt to provide students with the mathematical basis of chemistry and physics. Many of the subjects chosen are those that I wish that I had known when I was a student. It was just at that time that the no-mans-land between these two domains – chemistry and physics – was established by the "Harvard School", certainly attributable to E. Bright Wilson, Jr., J. H. van Vleck and the others of that epoch. I was most honored to have been a product, at least indirectly, of that group as a graduate student of J. C. Decius. Later, in my post-doc years, I profited from the Harvard–MIT seminars. During this experience I listened to, and tried to understand, the presentations by those most prestigious persons, who played a very important role in my development in chemistry and physics. The essential books at that time were most certainly the many publications by John C. Slater and the "Bible" on mathematical methods, by Margeneau and Murphy. They were my inspirations.

The expression "Chemical Physics" appears to have been coined by Slater. I should like to quote from the preface to his book, "*Introduction to Chemical Physics*" (McGraw-Hill, New York, 1939).

> *It is probably unfortunate that physics and chemistry ever were separated. Chemistry is the science of atoms and of the way in which they combine. Physics deals with the interatomic forces and with the large-scale properties of matter resulting from those forces. So long as chemistry was largely empirical and nonmathematical, and physics had not learned how to treat small-scale atomic forces, the two sciences seemed widely separated. But with statistical mechanics and the kinetic theory on the one hand and physical chemistry on the other, the two sciences began to come together. Now [1939!] that statistical mechanics has led to quantum theory and wave mechanics, with its explanations of atomic interactions, there is really nothing separating them any more . . ..*
>
> *A wide range of study is common to both subjects. The sooner we realize this the better. For want of a better name, as Physical Chemistry is already preempted, we may call this common field Chemical Physics. It is an overlapping field in which both physicists and chemists should*

> *be trained. There seems no valid reason why their training in it should differ . . .*

In the opinion of the present author, nobody could say it better.

That chemistry and physics are brought together by mathematics is the "raison d'être" of the present volume. The first three chapters are essentially a review of elementary calculus. After that there are three chapters devoted to differential equations and vector analysis. The remainder of the book is at a somewhat higher level. It is a presentation of group theory and some applications, approximation methods in quantum chemistry, integral transforms and numerical methods.

This is not a fundamental mathematics book, nor is it intended to serve a textbook for a specific course, but rather as a reference for students in chemistry and physics at all university levels. Although it is not computer-based, I have made many references to current applications – in particular to try to convince students that they should know more about what goes on behind the screen when they do one of their computer experiments. As an example, most students in the sciences now use a program for the fast Fourier transform. How many of them have any knowledge of the basic mathematics involved?

The lecture notes that I have written over many years in several countries have provided a basis for this book. More recently, I have distributed an early version to students at the third and fourth years at the University of Lille. It has been well received and found to be very useful. I hope that in its present form the book will be equally of value to students throughout their university studies.

The help of Professor Daniel Couturier, the ASA (Association de Solidarité des Anciens de l'Université des Sciences et Technologies de Lille) and the CRI (Centre de Resources Informatiques) in the preparation of this work is gratefully acknowledged. The many useful discussions of this project with Dr A. Idrissi, Dr F. Sokolić, Dr R. Withnall, Prof. M. Walters, Prof. D. W. Robinson and Prof. L. A. Veguillia-Berdicía are much appreciated.

My wife, Irène, and I have nicknamed this book "Mathieu". Throughout its preparation Irène has always provided encouragement – and patience when Mathieu was a bit trying or "Miss Mac" was in her more stubborn moods.

George Turrell
Lille, May 1, 2001

1 Variables and Functions

1.1 INTRODUCTION

The usual whole numbers, integers such as $1, 2, 3, 4 \ldots$, are usually referred to as Arabic numerals. It seems, however, that the basic decimal counting system was first developed in India, as it was demonstrated in an Indian astronomic calendar which dates from the third century AD. This system, which was composed of nine figures and the zero, was employed by the Arabs in the ninth century. The notation is basically that of the Arabic language and it was the Arabs who introduced the system in Europe at the beginning of the eleventh century.

In Europe the notion of the zero evolved slowly in various forms. Eventually, probably to express debts, it was found necessary to invent negative integers. The requirements of trade and commerce lead to the use of fractions, as ratios of whole numbers. However, it is obviously more convenient to express fractions in the form of decimals. The ensemble of whole numbers and fractions (as ratios of whole numbers) is referred to as rational numbers. The mathematical relation between decimal and rational fractions is of importance, particularly in modern computer applications.

As an example, consider the decimal fraction $x = 0.616161\cdots$. Multiplication by 100 yields the expression $100x = 61.6161\cdots = 61 + x$ and thus, $x = 61/99$, is a rational fraction. In general, if a decimal expression contains an infinitely repeating set of digits (61 in this example), it is a rational number. However, most decimal fractions do not contain a repeating set of digits and so are not rational numbers. Examples such as $\sqrt{3} = 1.732051\cdots$ and $\pi = 3.1415926536\cdots$ are irrational numbers.* Furthermore, the logarithms and trigonometric functions of most arguments are irrational numbers.

---

*A mnemonic for π based on the number of letters in words of the English language is quoted here from "the Green Book", Ian Mills, *et al.* (eds), "Quantities, Units and Symbols in Physical Chemistry", Blackwell Scientific Publications, London (1993):

> *'How I like a drink, alcoholic of course, after the heavy lectures involving quantum mechanics!'*

In practice, in numerical calculations with a computer, both rational and irrational numbers are represented by a finite number of digits. In both cases, then, approximations are made and the errors introduced in the result depend on the number of significant figures carried by the computer – the machine precision.* In the case of irrational numbers such errors cannot be avoided.

The ensemble of rational and irrational numbers are called real numbers. Clearly, the sum, difference and product of two real numbers is real. The division of two real numbers is defined in all cases but one – division by zero. Your computer will spit out an error message if you try to divide by zero!

1.2 FUNCTIONS

If two real variables are related such that, if a value of x is given, a value of y is determined, y is said to be a function of x. Thus, values may be assigned to x, the independent variable, leading to corresponding values of y, the dependent variable.

As an example, consider one mole of a gas at constant temperature. The volume V is a function of the applied pressure P. This relation can be expressed mathematically in the form

$$V = f(P), \tag{1}$$

or, $V = V(P)$. Note that to complete the functional relationship, the nature of the gas, as well as the temperature T, must be specified. A physical chemist should also insist that the system be in thermodynamic equilibrium.

In the case of an ideal gas, the functional relationship of Eq. (1) becomes

$$V = \frac{C(T)}{P}, \tag{2}$$

where $C(T)$ is a positive constant which is proportional to the absolute temperature T. Clearly, the roles of V and P can be reversed leading to the relation

$$P = \frac{C(T)}{V}. \tag{3}$$

The question as to which is the independent variable and which is the dependent one is determined by the way in which the measurements are made and, mathematically, on the presentation of the experimental data.

---

*Note that the zero is a special case, as its precision is not defined. Normally, the computer automatically uses the precision specified for other numbers.

Suppose that a series of measurements of the volume of a gas is made, as the applied pressure is varied. As an example, the original results obtained by Boyle* are presented as in Table 1.

In this case V is a function of P, but it is not continuous. It is the discrete function represented by the points shown in Fig. 1a. It is only the mathematical function of Eq. (2) that is continuous. If, from the experimental data, it is of interest to calculate values of V at intermediate points, it is necessary to estimate them with the use of, say, linear interpolation, or better, a curve-fitting procedure. In the latter case the continuous function represented by Eq. (2)

Table 1 Volume of a gas as a function of pressure.

V, Volume[a]	P, Pressure (inches Hg)	$1/V$
12	29.125	0.0833
10	35.3125	0.1
8	44.1875	0.125
6	58.8125	0.1667
5	70.6875	0.2
4	87.875	0.25
3	116.5625	0.3333

[a] Measured distance (inches) in a tube of constant diameter.

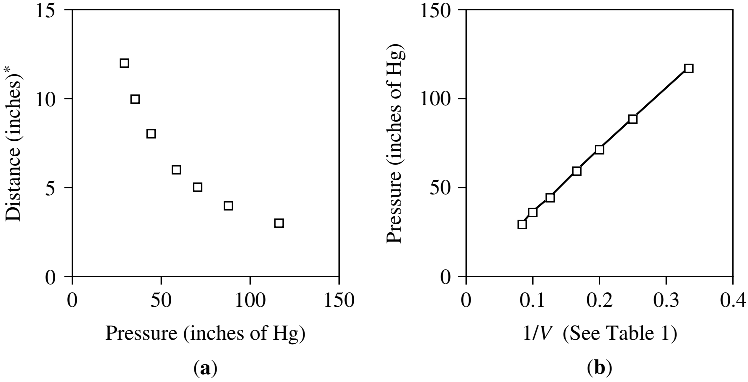


Fig. 1 Volume of a gas (expressed as distance in a tube of constant diameter) *versus* pressure (**a**). Pressure as a function of reciprocal volume (**b**).

---

*Robert Boyle, Irish physical chemist (1627–1691) would normally be employed. These questions, which concern the numerical treatment of data, will be considered in Chapter 13.

In Boyle's work the pressure was subsequently plotted as a function of the reciprocal of the volume, as calculated here in the third column of Table 1. The graph of P *vs.* $1/V$ is shown in Fig. 1b. This result provided convincing evidence of the relation given by Eq. (3), the mathematical statement of Boyle's law. Clearly, the slope of the straight line given in Fig. 1b yields a value of $C(T)$ at the temperature of the measurements [Eq. (3)] and hence a value of the gas constant R. However, the significance of the temperature was not understood at the time of Boyle's observations.

In many cases a series of experimental results are not associated with a known mathematical function. In the following example Miss X weighed herself each morning beginning on the first of February. These data are presented graphically as shown in Fig. 2. Here interpolated points are of no significance, nor is extrapolation. By extrapolation Miss X would weigh nearly nothing in a year-and-a-half or so. However, as the data do exhibit a trend over a relatively short time, it is useful to employ a curve-fitting procedure. In this example Miss X might be happy to conclude that on the average she lost 0.83 kg per week during this period, as indicated by the slope of the straight line in Fig. 2.

Now reconsider the function given by Eq. (3). It has the form of a hyperbola, as shown in Fig. 3. Different values of $C(T)$ lead to other members of the family of curves shown. It should be noted that this function is antisymmetric with respect to the inversion operation $V \to -V$ (see Chapter 8). Thus, P is said to be an odd function of V, as $P(V) = -P(-V)$.

It should be evident that the negative branches of P *vs.* V shown in Fig. 3 can be excluded. These branches of the function are correct mathematically,

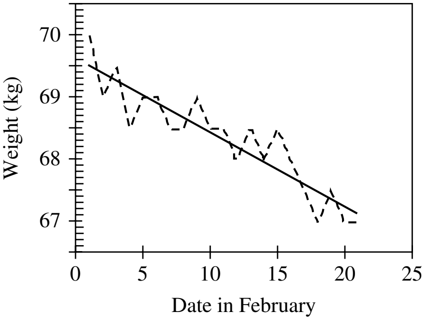


Fig. 2 Miss X's weight as a function of the date in February. The straight line is obtained by a least-squares fit to the experimental data (see Chapter 13).

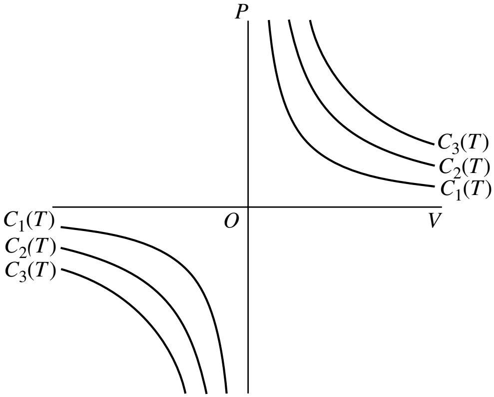


Fig. 3 Pressure *versus* volume [Eq. (3)], with $C_3(T) > C_2(T) > C_1(T)$.

but are of no physical significance for this problem. This example illustrates the fact that functions may often be limited to a certain domain of acceptability. Finally, it should be noted that the function $P(V)$ presented in Fig. 3 is not continuous at the origin ($V = 0$). Therefore, from a physical point of view the function is only significant in the region $0 < V < \infty$. Furthermore, physical chemists know that Eqs. (2) and (3) do not apply at high pressures because the gas is no longer ideal.

As $C(T)$ is a positive quantity, Eq. (3) can be written in the form

$$ln\ P = ln\ C - ln\ V. \tag{4}$$

Clearly a plot of $ln\ P$ *vs.* $ln\ V$ at a given (constant) temperature yields a straight line with an intercept equal to $ln\ C$. This analysis provides a convenient graphical method of determining the constant C.

It is often useful to shift the origin of a given graph. Thus, for the example given above consider that the axes of V and P are displaced by the amounts v and p, respectively. Then, Eq. (3) becomes

$$P - p = \frac{C(T)}{V - v}, \tag{5}$$

and the result is as plotted in Fig. 4. The general hyperbolic form of the curves has not been changed, although the resulting function $P(V)$ is no longer odd – nor is it even.

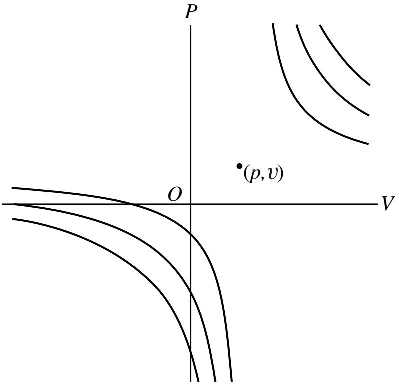


Fig. 4 Plots of Eq. (5) for given values of v and p.

1.3 CLASSIFICATION AND PROPERTIES OF FUNCTIONS

Functions can be classified as either algebraic or transcendental. Algebraic functions are rational integral functions or polynomials, rational fractions or quotients of polynomials, and irrational functions. Some of the simplest in the last category are those formed from rational functions by the extraction of roots. The more elementary transcendental functions are exponentials, logarithms, trigonometric and inverse trigonometric functions. Examples of these functions will be discussed in the following sections.

When the relation $y = f(x)$ is such that there is only one value of y for each acceptable value of x, $f(x)$ is said to be a single-valued function of x. Thus, if the function is defined for, say, $x = x_1$, the vertical line $x = x_1$ intercepts the curve at one and only one point, as shown in Fig. 5. However, in many cases a given value of x determines two or more distinct values of y.

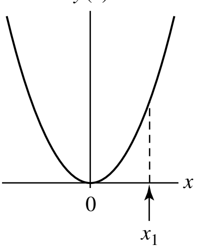


Fig. 5 Plot of $y = x^2$.

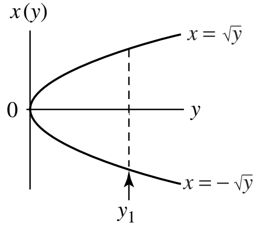


Fig. 6 Plot of $x = \pm\sqrt{y}$.

The curve shown in Fig. 5 can be represented by

$$y = x^2, \tag{6}$$

where y has the form of the potential function for a harmonic oscillator (see Chapter 5). This function is an even function of x, as $y(x) = y(-x)$. Clearly, y is a single-valued function of x. Now, if Eq. (6) is rewritten in the equivalent form

$$x^2 = y, \quad y \geq 0 \tag{7}$$

it defines a double-valued function whose branches are given by $x = \sqrt{y}$ and $x = -\sqrt{y}$. These branches are the upper and lower halves of the parabola shown in Fig. 6. It should be evident from this example that to obtain a given value of x, it is essential to specify the particular branch of the (in general) multiple-valued function involved. This problem is particularly important in numerical applications, as carried out on a computer (Don't let the computer choose the wrong branch!).

1.4 EXPONENTIAL AND LOGARITHMIC FUNCTIONS

If $y = f(x)$ is given by

$$y = a^x \tag{8}$$

$y = f(x)$ is an exponential function. The independent variable x is said to be the argument of f. The inverse relation, the logarithm, can then be defined by

$$x = log_a y \tag{9}$$

and a is called the base of the logarithm. It is clear, then, that $log_a a = 1$ and $log_a 1 = 0$. The logarithm is a function that can take on different values depending on the base chosen. If $a = 10$, log_{10} is usually written simply as log. A special case, which is certainly the most important in physics and chemistry, as well as in pure mathematics, is that with $a = e$. The quantity e, which serves as the base of the natural or Naperian* logarithm, $log_e \equiv ln$, can be defined by the series†

$$y = e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}. \tag{10}$$

It should be noted that the derivative of Eq. (10), taken term by term, is given by

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 0 + 1 + x + \frac{1}{2!}x^2 + \cdots = e^x. \tag{11}$$

Thus, $(\mathrm{d}/\mathrm{d}x)e^x = e^x$ and here the operator $\mathrm{d}/\mathrm{d}x$ plays the role of the identity with respect to the function $y = e^x$.‡ It will be employed in the solution of differential equations in Chapter 5.

Consider, now, the function $f(n) = (1 + 1/n)^n$. It is evaluated in Table 2 as a function of n, where it is seen that it approaches the value of $e \equiv lim_{n\to\infty}(1 + 1/n)^n = 2.7182818285\cdots$, an irrational number, as n becomes infinite. For simplicity, it has been assumed here that n is an integer, although it can be shown that the same limiting value is obtained for noninteger values of n. The identification of e with that employed in Eq. (10) can be made by

Table 2 Evaluation of $e \equiv lim_{n\to\infty}(1 + 1/n)^n$.

n	$f(n) = (1 + 1/n)^n$
1	2.000
2	2.250
5	2.489
10	2.594
20	2.653
50	2.691
100	2.705
1000	2.717
10 000	2.718
∞	2.7182818285

---

*John Napier or Neper, Scottish mathematician (1550–1617).

†The factorial $n! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot \cdots \cdot n$ (with $0! = 1$) has been introduced in Eq. (10). See also Section 4.5.4.

‡Note that e^x is often written $exp\,x$.

application of the binomial theorem (see Section 2.10). The functions e^x and $ln_e\, x$ are illustrated in Figs. 7 and 8, respectively.

As indicated above, the two logarithmic functions *ln* and *log* differ in the base used. Thus, if $y = e^x = 10^z$,

$$z = log\ y \tag{12}$$

and

$$x = ln\ y. \tag{13}$$

Then,

$$ln\ y = ln(e^x) = ln(10^z) \tag{14}$$

and, as $ln\ e = 1$ and $ln\ 10 = 2.303$,

$$ln\ y = 2.303\, log\ y. \tag{15}$$

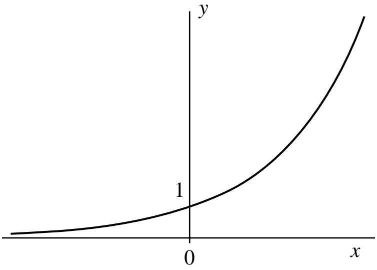


Fig. 7 Plot of $y(x) = e^x$.

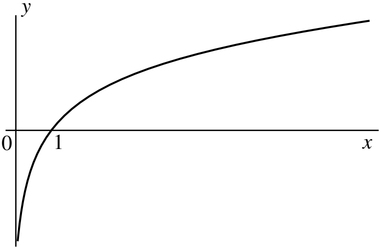


Fig. 8 Plot of $y(x) = ln_e\, x$.

The numerical factor 2.303 (or its reciprocal) appears in many formulas of physical chemistry and has often been the origin of errors in published scientific work. It is evident that these two logarithmic functions, *ln* and *log*, must be carefully distinguished.

It was shown above that the derivative of e^x is equal to e^x. Thus, if $x = ln\, y$, $y = e^x$ and

$$\frac{\mathrm{d}y}{\mathrm{d}x} = e^x = y. \tag{16}$$

Then,

$$\frac{\mathrm{d}x}{\mathrm{d}y} = \frac{\mathrm{d}}{\mathrm{d}y}(ln\, y) = \frac{1}{y} \tag{17}$$

and, as $\mathrm{d}x = \mathrm{d}y/y$,

$$x = \int \frac{1}{y}\mathrm{d}y = ln\, y + C, \tag{18}$$

where C is here the constant of integration (see Chapter 3).

1.5 APPLICATIONS OF EXPONENTIAL AND LOGARITHMIC FUNCTIONS

As an example of the use of the exponential and logarithmic functions in physical chemistry, consider a first-order chemical reaction, such as a radioactive decay. It follows the rate law

$$-\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} = k[\mathrm{A}], \tag{19}$$

where [A] represents the concentration of reactant A at time t. With the use of Eq. (18) this expression can be integrated to yield

$$-\,ln[\mathrm{A}] = kt + C, \tag{20}$$

where C is a constant.* The integration constant C can only be evaluated if additional data are available. Usually the experimentalist measures at a given time, say t_0, the concentration of reactant, $[\mathrm{A}]_0$. This relation, which constitutes an initial condition on the differential equation, Eq. (19), allows the integration constant C to be evaluated. Thus, $[\mathrm{A}] = [\mathrm{A}]_0$ at $t = t_0$, and

$$ln\,\frac{[\mathrm{A}]}{[\mathrm{A}]_0} = -kt. \tag{21}$$

---

*The indefinite integral is discussed in Section 3.1.

This expression can of course be written in the exponential form, *viz*.,

$$[\mathrm{A}] = [\mathrm{A}]_0 e^{-kt}, \tag{22}$$

the result that is plotted in Fig. 9.

In the case of radio-active decay the rate is often expressed by the half-life, namely, the time required for half of the reactant to disappear. From Eq. (22) the half-life is given by $t_{1/2} = (ln\, 2)/k$.

As a second example, consider the absorption of light by a thin slice of a given sample, as shown in Fig. 10. The intensity of the light incident on the sample is represented by I_0, while I is the intensity at a distance x. Following Lambert's law,* the decrease in intensity is given by

$$-\mathrm{d}I = \alpha I \mathrm{d}x, \tag{23}$$

where α is a constant. Integration of Eq. (23) leads to

$$-\, ln\, I = \alpha x + C. \tag{24}$$

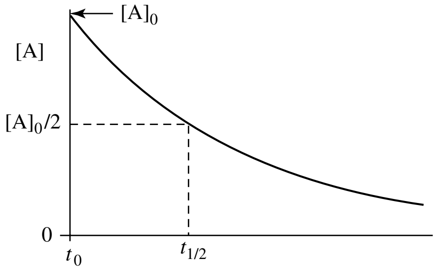


Fig. 9 Exponential decay of reactant in a first-order reaction.

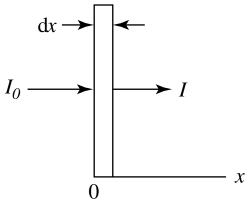


Fig. 10 Transmission of light through a thin slice of sample.

---

*Jean-Henri Lambert, French mathematician (1728–1777).

Here again, certain conditions must be imposed on the general solution of Eq. (24) to evaluate the constant of integration. They are in this case referred to as the boundary conditions. Thus if $I = I_0$ at $x = 0$, $C = -\, ln\, I_0$ and Eq. (24) becomes

$$ln\, \frac{I}{I_0} = -\alpha x. \tag{25}$$

For a sample of thickness $x = \ell$, the fraction of light transmitted is given by

$$\frac{I}{I_0} = e^{-\alpha\ell}. \tag{26}$$

The integration of Eq. (23) can also be carried out between limits (see Chapter 3), in the form

$$-\int_{I_0}^{I} \frac{\mathrm{d}I}{I} = \alpha \int_0^{\ell} \mathrm{d}x. \tag{27}$$

The result is then

$$ln\, \frac{I}{I_0} = -\alpha\ell \tag{28}$$

or

$$log\, \frac{I}{I_0} = -\left(\frac{\alpha}{2.303}\right)\ell. \tag{29}$$

Thus, the extinction coefficient, as usually defined in analytical spectroscopy, includes the factor 2.303 in the denominator. It should be apparent that the light intensity decreases exponentially within the sample, by analogy with the decrease in reactant in the previous example [Fig. (9)].

1.6 COMPLEX NUMBERS

Consider the relation $z = x + iy$, where x and y are real numbers and i has the property that $i^2 = -1$. The variable z is called a complex number, with real part x and imaginary part y. Thus, $\Re e[z] = x$ and $\Im m[z] = y$. It will be shown in Chapter 8 that the quantities $i^0 = 1, i^1 = i, i^2 = -1$ and $i^3 = -i$ form a group, a cyclic group of order four.

Two complex numbers which differ only in the sign of their imaginary parts are called complex conjugates – or simply conjugates. Thus, if $z = x + iy$, $z^\star = x - iy$ is its complex conjugate, which is obtained by replacing i by $-i$. Students are usually introduced to complex numbers as solutions to certain quadratic equations, where the roots always appear as conjugate pairs. It should be noted that in terms of absolute values $|z| = |z^{\star}| = \sqrt{x^2 + y^2}$, which is sometimes called the modulus of z.

It is often convenient to represent complex numbers graphically in what is referred to as the complex plane.* The real numbers lie along the x axis and the pure imaginaries along the y axis. Thus, a complex number such as $3 + 4i$ is represented by the point (3,4) and the locus of points for a constant value of $r = |z|$ is a circle of radius $|z|$ centered at the origin, as shown in Fig. 11. Clearly, $x = r\,cos\,\varphi$ and $y = r\,sin\,\varphi$, and in polar coordinates

$$z = x + iy = re^{i\varphi} \tag{30}$$

$$= r(cos\,\varphi + i\,sin\,\varphi). \tag{31}$$

Then,

$$e^{i\varphi} = cos\,\varphi + i\,sin\,\varphi, \tag{32}$$

which is the very important relation known as Euler's equation.† It should be emphasized here that the exponential functions of both imaginary and real arguments are of extreme importance. They will be discussed in some detail in Chapter 11 in connection with the Fourier and Laplace transforms, respectively.

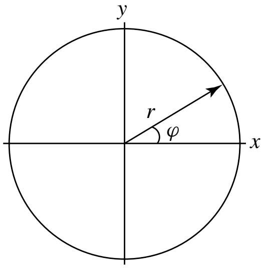


Fig. 11 Circle of radius $r = |z|$ in the complex plane.

---

*This system of representing complex numbers was developed by Jean-Robert Argand, Swiss mathematician (1768–1822), among others, near the beginning of the 19th century.

†Leonhard Euler, Swiss mathematician (1707–1783). This relation is sometimes attributed to Abraham De Moivre, British mathematician (1667–1754).

1.7 CIRCULAR TRIGONOMETRIC FUNCTIONS

The exponential function was defined in Eq. (10) terms of an infinite series. By analogy, the left-hand side of Eq. (32) can be expressed in the form

$$e^{i\varphi} = 1 + \frac{i\varphi}{1!} + \frac{(i\varphi)^2}{2!} + \cdots + \frac{(i\varphi)^n}{n!} + \cdots, \tag{33}$$

which can be separated into its real and imaginary parts, *viz.*

$$\Re e[e^{i\varphi}] = 1 - \frac{\varphi^2}{2!} + \frac{\varphi^4}{5!} - \cdots = cos\ \varphi \tag{34}$$

and

$$\Im m[e^{i\varphi}] = \varphi - \frac{\varphi^3}{3!} + \frac{\varphi^5}{5!} - \cdots = sin\ \varphi. \tag{35}$$

Comparison with Eq. (32) yields the last equalities in Eqs. (34) and (35). The infinite series in these two equations are often taken as the fundamental definitions of the cosine and sine functions, respectively. The equivalent expressions for these functions,

$$cos\ \varphi = \frac{e^{i\varphi} + e^{-i\varphi}}{2} \tag{36}$$

and

$$sin\ \varphi = \frac{e^{i\varphi} - e^{-i\varphi}}{2i} \tag{37}$$

can be easily derived from Eqs. (33–35). Alternatively, they can be used as definitions of these functions. The functions $cos\ \varphi$ and $sin\ \varphi$ are plotted *versus* φ expressed in radians in Figs. 12a and 12b, respectively. The two curves have the same general form, with a period of 2π, although they are "out of phase" by $\pi/2$. It should be noted that the functions cosine and sine are even and odd functions, respectively, of their arguments.

In some applications it is of interest to plot the absolute values of the cosine and sine functions in polar coordinates. These graphs are shown as Figs. 13a and 13b, respectively.

From Eqs. (36) and (37) it is not difficult to derive the well-known relation

$$sin^2\ \varphi + cos^2\ \varphi = 1, \tag{38}$$

which is applicable for all values of φ. Dividing each term by $cos^2\ \varphi$ leads to the expression

$$\frac{sin^2\ \varphi}{cos^2\ \varphi} + 1 = \frac{1}{cos^2\ \varphi} \tag{39}$$

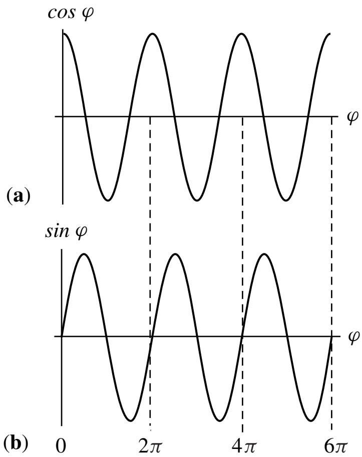
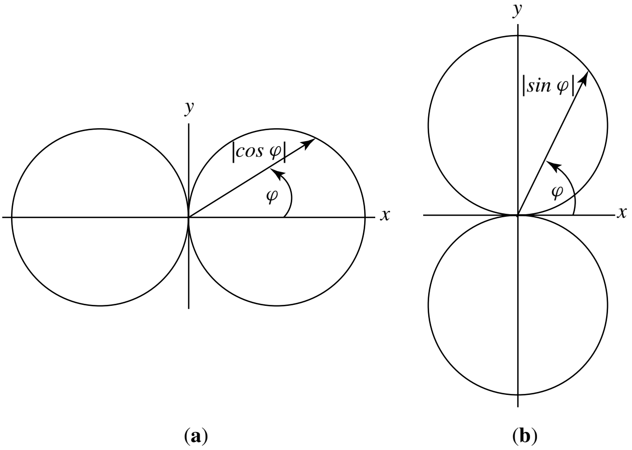


Fig. 12 The functions (**a**) cosine and (**b**) sine.



(**a**) (**b**)

Fig. 13 Absolute values of (**a**) cosine and (**b**) sine in polar coordinates.

or,

$$tan^2 \varphi + 1 = sec^2 \varphi. \tag{40}$$

Similarly, it is easily found that

$$1 + \frac{cos^2 \varphi}{sin^2 \varphi} = \frac{1}{sin^2 \varphi} \tag{41}$$

and thus

$$1 + cot^2\,\varphi = csc^2\,\varphi. \tag{42}$$

Equations (40) and (42) define the trigonometric functions tangent (*tan*), cotangent (*cot*), secant (*sec*) and cosecant (*csc*). These relations were probably learned in high school, but are in any case available on most present-day calculators. The various formulas for more complicated arguments of the trigonometric functions can all be derived from the definitions given in Eqs. (36) and (37). For example, the relation involving the arguments α and β,

$$cos(\alpha - \beta) = cos\,\alpha\,cos\,\beta + sin\,\alpha\,sin\,\beta \tag{43}$$

can be obtained without too much difficulty.

1.8 HYPERBOLIC FUNCTIONS

The trigonometric functions developed in the previous section are referred to as circular functions, as they are related to the circle shown in Fig. 11. Another somewhat less familiar family of functions, the hyperbolic functions, can also be derived from the exponential. They are analogous to the circular functions considered above and can be defined by the relations

$$cosh\,\varphi = \frac{e^{\varphi} + e^{-\varphi}}{2} \tag{44}$$

and

$$sinh\,\varphi = \frac{e^{\varphi} - e^{-\varphi}}{2}. \tag{45}$$

The first of these functions is effectively the sum of two simple exponentials, as shown in Fig. 14a, while the hyperbolic sine (*sinh*) is the difference [Eq. (44) and Fig. 14b]. It should be noted that the hyperbolic functions have no real period. They are periodic in the imaginary argument $2\pi i$.

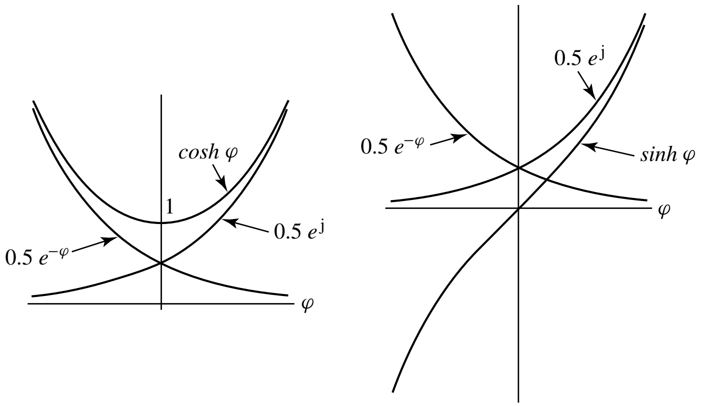
The hyperbolic and circular functions are related *via* the expressions

$$cosh\,\varphi = cos\,i\varphi, \quad cos\,\varphi = cosh\,i\varphi \tag{46}$$

and

$$sinh\,\varphi = \frac{1}{i}\,sin\,i\varphi, \quad sin\,\varphi = \frac{1}{i}\,sinh\,i\varphi. \tag{47}$$

Because of this duality, every relation involving circular functions has its formal counterpart in the corresponding hyperbolic functions, and *vice versa.*



(**a**) (**b**)

Fig. 14 The hyperbolic functions (**a**) *cosh* and (**b**) *sinh*.

Thus, the various relations between the hyperbolic functions can be derived as carried out above for the circular functions. For example,

$$cosh^2\varphi - sinh^2\,\varphi = 1, \tag{48}$$

which is analogous to Eq. (38), as illustrated by problem 21.

PROBLEMS

1. Given the decimal 0.7171 . . . , find two numbers whose ratio yields the same value. Ans. 71, 99

2. Repeat question 1 for the decimal 18.35912691269126 . . . . Ans. 18357291, 999900

3. The fraction 22/7 is often used to approximate the value of π. Calculate the error resulting from the use of this approximation. Ans. 0.04%

4. Calculate the values of the expression $\sqrt{log_3 3}$. Ans. ±1

5. Calculate the values of the expressions $log\ 10^{-3}$ and $ln\ 10^{-3}$. Ans. −3, −6.909

6. Calculate the value of the constant a for which the curve $y = ln((5-x)/(8-x)a)$ passes through the point (1,1). Ans. $a = 7e/4$

7. Derive the general relation between the temperature expressed in degrees Fahrenheit* (F) and degrees Celsius† (C). Ans. $F = \frac{9}{5}C + 32$

8. The length ℓ of an iron bar varies linearly with the temperature over a certain range. At 15°C its length is 1 m. Its length increases by 12 μm/°C. Derive the general relation for ℓ as a function of the temperature t.
Ans. $\ell = 12 \times 10^{-6}t + 0.99982$

9. Calculate the rate constant for a first-order chemical reaction which is 90% completed in 10 min [see Eq. (21)]. Ans. 0.23 min^{-1}

10. A laser beam was used to measure light absorption by a bottle of Bordeaux (1988). In the middle of the bottle (diameter D) 60% of the light was absorbed. At the neck of the bottle (diameter d) it was only 27%. Calculate the ratio of the diameters of the bottle, D/d. What approximations were made in this analysis?
Ans. 2.91

11. With a complex number z defined by Eqs. (30) and (31), find an expression for z^{-1}. Ans. $(1/r)(cos\,\varphi - i\,sin\,\varphi)$

12. Find all of the roots of $\sqrt[4]{16}$. Ans. $2, -2, 2i, -2i$

13. Find all of the roots of the equation $x^3 + 27 = 0$. Ans. $-3, \frac{3}{2}(1 \pm i\sqrt{3})$

14. Given $e^x - e^{-x} = 1, e^x > 1$, find x. Ans. $ln[(1 + \sqrt{5})/2]$

15. Derive the expression for $x(y)$, where $y = ln(e^{2x} - 1)$. Ans. $x = ln\sqrt{e^y + 1}$

16. Write the function $(i + 3)/(i - 1)$ in the form $a + bi$, where $i \equiv \sqrt{-1}$ and a and b are real. Ans. $-1 - 2i$

17. Repeat question 16 for the function $((3i - 7)/(i + 4))$.
Ans. $-(25/17) + (19/17)i$

18. Find the absolute value of the function $(2i - 1)/(i - 2)$. Ans. 1

19. Repeat problem 18 for the function $3i/(i - \sqrt{3})$. Ans. 3/2

20. Given the definitions $cos\,\varphi = (e^{i\varphi} + e^{-i\varphi})/2$ and, $sin\,\varphi = (e^{i\varphi} - e^{-i\varphi})/2i$, show that $cos(\varphi + \gamma) = cos\,\varphi\,cos\,\gamma - sin\,\varphi\,sin\,\gamma$ and therefore, $cos[(\pi/2) - \varphi] = sin\,\varphi$.

21. Given the definitions of the functions *sinh* and *cosh*, prove Eq. (48).

22. Show that $sinh^{-1}x = ln(x + \sqrt{x^2 + 1}), \quad x > 0$.

---

*Daniel Gabriel Fahrenheit, German physicist (1686–1736).

†Anders Celsius, Swedish astronomer and physicist (1701–1744).

2 Limits, Derivatives and Series

2.1 DEFINITION OF A LIMIT

Given a function $y = f(x)$ and a constant a: If there is a number, say γ, such that the value of $f(x)$ is as close to γ as desired, where x is different from a, then the limit of $f(x)$ as x approaches a is equal to γ. This formalism is then written as,

$$\lim_{x \to a} f(x) = \gamma. \tag{1}$$

A graphical interpretation of this concept is shown in Fig. 1.

If there is a value of ε such that $|f(x) - \gamma| < \varepsilon$, then x can be chosen anywhere at a value δ from the point $x = a$, with $0 < |x - a| < \delta$. Thus it is possible in the region near $x = a$ on the curve shown in Fig. 1, to limit the variation in $f(x)$ to as little as desired by simply narrowing the vertical band around $x = a$. Thus, Eq. (1) is graphically demonstrated. It should be emphasized that the existence of the limit given by Eq. (1) does not necessarily mean that $f(a)$ is defined.

As an example, consider the function

$$y(x) = \frac{\sin x}{x}. \tag{2}$$

The function *sin x* can be defined by an infinite series, as given in Eq. (1-35). Division by x yields the series

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \cdots. \tag{3}$$

It is evident from the right-hand side of Eq. (3) that this function becomes equal to 1 as x approaches zero, even though $y(0) = \frac{0}{0}$.* Thus, from a mathematical point of view it is not continuous, as it is not defined at $x = 0$. This function, which is of extreme importance in the applications of the Fourier transform (Chapter 11), is presented in Fig. 2.

---

*This result, $\frac{0}{0}$, is the most common indeterminate form (see Section 2.8).

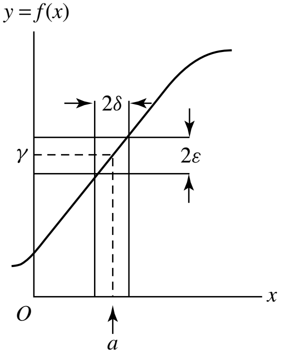


Fig. 1 The limit of a function.

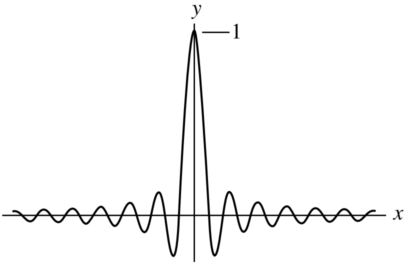


Fig. 2 The function $y(x) = \dfrac{\sin x}{x}$.

It should be noted that computer programs written to calculate $y(x) = \sin x/x$ will usually fail at the point $x = 0$. The computer will display a "division by zero" error message. The point $x = 0$ must be treated separately and the value of the limit ($y = 1$) inserted. However, "intelligent" programs such as *Mathematica** avoid this problem.

It is often convenient to consider the limiting process described above in the case of a function such as shown in Fig. 3. Then, it is apparent that the limiting value of $f(x)$ as $x \to a$ depends on the direction chosen. As x approaches a from the left, that is, from the region where $x < a$,

$$\lim_{x \to a_-} f(x) = \gamma_- . \tag{4}$$

**Mathematica*, Wolfram Research, Inc., Champaign, Ill., 1997.

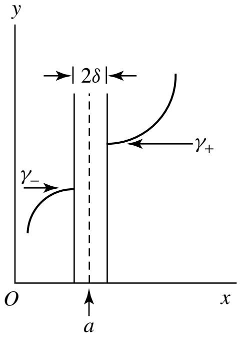


Fig. 3 The limits of the function $f(x)$ as x approaches a.

Similarly, from the right the limit is given by

$$\lim_{x \to a_+} f(x) = \gamma_+. \tag{5}$$

Clearly, in this example the two limits are not the same and this function cannot be evaluated at $x = a$. Another example is that shown in Fig. (1-3), where $P \to \infty$ as V approaches zero from the right (and $-\infty$, if the approach were made from the left).

2.2 CONTINUITY

The notion of continuity was introduced in Chapter 1. However, it can now be defined more specifically in terms of the appropriate limits.

A function $f(x)$ is said to be continuous at the point $x = a$ if the following three conditions are satisfied:

(i) The function is defined at $x = a$, namely, $f(a)$ exists,
(ii) The function approaches a limit as x approaches a (in either direction), *i.e.* $\lim_{x \to a} f(x)$ exists and
(iii) The limit is equal to the value of the function at the point in question, *i.e.* $\lim_{x \to a} f(x) = f(a)$.

See problem 3 for some applications.

The rules for combining limits are, for the most part, obvious:

(i) The limit of a sum is equal to the sum of the limits of the terms; thus, $\lim_{x \to a}[f(x) + g(x)] = \lim_{x \to a} f(x) + \lim_{x \to a} g(x)$.

(ii) The limit of a product is equal to the product of the limits of the factors; then $lim_{x\to a}[f(x)\cdot g(x)] = lim_{x\to a} f(x)\cdot lim_{x\to a} g(x)$ and hence $lim_{x\to a}[f(cx)] = c\,lim_{x\to a} f(x)$, where c is an arbitrary constant.

(iii) The limit of the quotient of two functions is equal to the quotient of the limits of the numerator and denominator, if the limit of the denominator is different from zero, *viz.*

$$\lim_{x\to a}\frac{f(x)}{g(x)} = \frac{\lim\limits_{x\to a} f(x)}{\lim\limits_{x\to a} g(x)},\quad \text{if } \lim_{x\to a} g(x)\neq 0.$$

Rule (iii) is particularly important in the tests for series convergence that will be described in Section 2.11.

An additional question arises in the application of rule (iii) when both the numerator and the denominator approach zero. This rule does not then apply; the ratio of the limits becomes in this case $\frac{0}{0}$, which is undefined. However, the limit of the ratio may exist, as found often in the applications considered in the following chapters. In fact, an example has already been presented [see Eq. (2)].

2.3 THE DERIVATIVE

Given a continuous function $y = f(x)$, for a given value of x there is a corresponding value of y. Now, consider another value of x which differs from the first one by an amount Δx, which is referred to as the increment of x. For this value of x, y will have a different value which differs from the first one by a quantity Δy. Thus,

$$y + \Delta y = f(x + \Delta x) \tag{6}$$

or

$$\Delta y = f(x + \Delta x) - f(x) \tag{7}$$

and

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}. \tag{8}$$

In the limit as both the numerator and the denominator of Eq. (8) approach zero, the finite differences Δy and Δx become the (infinitesimal) differentials dy and dx. Thus, Eq. (8) takes the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \lim_{\Delta x\to 0}\frac{f(x + \Delta x) - f(x)}{\Delta x}, \tag{9}$$

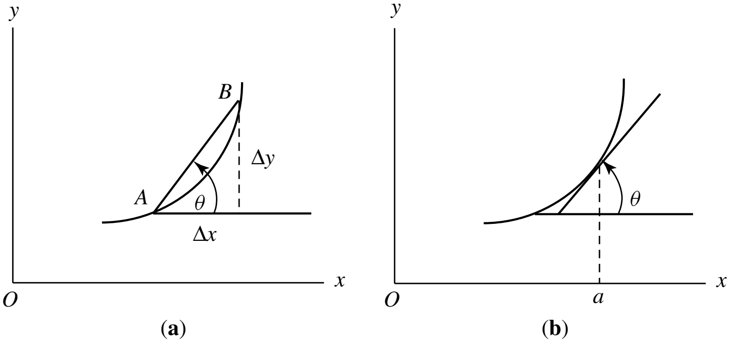


Fig. 4 The derivative: (**a**) Definitions of Δx and Δy; (**b**) $tan\,\theta = \dfrac{\mathrm{d}y}{\mathrm{d}x}$.

which is called the derivative of y with respect to x. The notation $y' \equiv \mathrm{d}y/\mathrm{d}x$ is often used if there is no ambiguity regarding the independent variable x. The derivative exists for most continuous functions. As shown in elementary calculus, the requirements for the existence of the derivative in some range of values of the independent variable, are that it be continuous, single-valued and differentiable, that is, that y be an analytic function of x.

A graphical interpretation of the derivative is introduced here, as it is extremely important in practical applications. The quantities Δx and Δy are identified in Fig. 4a. It should be obvious that the ratio, as given by Eq. (8) represents the tangent of the angle θ and that in the limit (Fig. 4b), the slope of the line segment $\overline{AB}$ (the secant) becomes equal to the derivative given by Eq. (8).

It was already assumed in Chapter 1 that readers are familiar with the methods for determining the derivatives of algebraic functions. The general rules, as proven in all basic calculus courses, can be summarized as follows.

(i) Derivative of a constant:

$$\frac{\mathrm{d}a}{\mathrm{d}x} = 0, \tag{10}$$

where a is a constant.

(ii) Derivative of a sum:

$$\frac{\mathrm{d}}{\mathrm{d}x}(u + v) = \frac{\mathrm{d}u}{\mathrm{d}x} + \frac{\mathrm{d}v}{\mathrm{d}x}, \tag{11}$$

where u and v are functions of x.

(iii) Derivative of a product:

$$\frac{\mathrm{d}}{\mathrm{d}x}(uv) = u\frac{\mathrm{d}v}{\mathrm{d}x} + v\frac{\mathrm{d}u}{\mathrm{d}x}. \tag{12}$$

(iv) Derivative of a quotient:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{u}{v}\right) = \frac{v\dfrac{\mathrm{d}u}{\mathrm{d}x} - u\dfrac{\mathrm{d}v}{\mathrm{d}x}}{v^2}. \tag{13}$$

(v) Derivative of a function of a function:
Given the function $y[u(x)]$,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u}\cdot\frac{\mathrm{d}u}{\mathrm{d}x}. \tag{14}$$

Equation (14) leads immediately to the relations

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\dfrac{\mathrm{d}y}{\mathrm{d}u}}{\dfrac{\mathrm{d}x}{\mathrm{d}u}} \text{ if } \frac{\mathrm{d}x}{\mathrm{d}u} \neq 0$$

and

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\dfrac{\mathrm{d}x}{\mathrm{d}y}} \text{ if } \frac{\mathrm{d}x}{\mathrm{d}y} \neq 0.$$

(vi) The power formula:

$$\frac{\mathrm{d}}{\mathrm{d}x}u^n = nu^{n-1}\frac{\mathrm{d}u}{\mathrm{d}x} \tag{15}$$

for the function $u(x)$ raised to any power.

The derivative of the logarithm was already discussed in Chapter 1, while the derivatives of the various trigonometric functions can be developed from their definitions [see, for example, Eqs. (1-36), (1-37), (1-44) and (1-45)]. A number of expressions for the derivatives can be derived from the problems at the end of this chapter.

2.4 HIGHER DERIVATIVES

If y is a function of x, the derivative of $y(x)$ is also, in general, a function of x. It can then be differentiated to yield the second derivative of y with respect to x, namely,

$$y'' \equiv \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\mathrm{d}y}{\mathrm{d}x}. \tag{16}$$

It should be noted here that the operation of taking the derivative, that is, the result of the operator $\mathrm{d}/\mathrm{d}x$ operating on a function of x, followed by the same operation, yields the second derivative. Thus, the successive application of two operators is referred to as their product. This question is addressed more specifically in Chapter 7.

Clearly, y'' in Eq. (16) represents the rate of change of the slope of the function $y(x)$. The second derivative can be expressed in terms of derivatives with respect to y, *viz.*,

$$\frac{\mathrm{d}y'}{\mathrm{d}x} = \frac{\dfrac{\mathrm{d}y'}{\mathrm{d}y}}{\dfrac{\mathrm{d}x}{\mathrm{d}y}}, \tag{17}$$

which leads to

$$y'' \equiv \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = -\frac{\dfrac{\mathrm{d}^2 x}{\mathrm{d}y^2}}{\left(\dfrac{\mathrm{d}x}{\mathrm{d}y}\right)^3}, \tag{18}$$

a relation which is sometimes useful.

2.5 IMPLICIT AND PARAMETRIC RELATIONS

Often two variables x and y are related implicitly in the form $f(x, y) = 0$. Although it is sometimes feasible to solve for y as a function of x, such is not always the case. However, if rule (vi) above [Eq. (15)] is applied with care, the derivatives can be evaluated. As an example, consider the equation for a circle of radius r,

$$x^2 + y^2 = r^2. \tag{19}$$

Rather than to solve for y, it is more convenient to apply rule (vi) directly; then $2x + 2y(\mathrm{d}y/\mathrm{d}x) = 0$ and $y' = \mathrm{d}y/\mathrm{d}x = -x/y$. The second derivative is then obtained with the use of rule (iv):

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \frac{xy' - y}{y^2}.$$

Sometimes the two variables are expressed in terms of a third variable, or parameter. Then, $x = u(t)$ and $y = v(t)$ and, in principle, the parameter t can be eliminated to obtain a relation between x and y. Here again, this operation is not always easy, or even, possible. An example is provided by the pair of equations $x = t^2 + 2t - 4$ and $y = t^2 - t + 2$. The two derivatives $\mathrm{d}x/\mathrm{d}t$ and $\mathrm{d}y/\mathrm{d}t$ are easily obtained and, with the aid of rule (v), their ratio becomes

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{3t^2 - 1}{3t^2 + 2}.$$

The second derivative is then given by

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{18t}{(3t^2 + 2)^3},$$

where the relations below Eq. (14) have been employed. This result is left as an exercise for the reader (problem 7).

2.6 THE EXTREMA OF A FUNCTION AND ITS CRITICAL POINTS

As shown in Fig. 4b, the derivative of a function evaluated at a given point is equal to the slope of the curve at that point. Given two points x_1 and x_2 in the neighborhood of a such that $x_1 < a$ and $x_2 > a$, it is apparent that if $f(x_1) < f(a) < f(x_2)$, the slope is positive. Similarly, if $f(x_1) > f(a) > f(x_2)$ in the same region, the slope is negative. On the other hand, if $f(x_1) < f(a) > f(x_2)$ the function has a maximum value in the neighborhood of a. It is of course minimal in that region if $f(x_1) > f(a) < f(x_2)$. At either a maximum or a minimum the derivative of the function is zero. Thus, the slope is equal to zero at these points, which are the *extrema*, as shown by points a and c in Fig. 5. A function may have additional maxima or minima in other regions. In Fig. 5 there are maxima at $x = a$ and $x = b$. As $f(a) > f(b)$, the point a is called the absolute or principal maximum and that at b is a submaximum.

It should be obvious from Fig. 5 that the curve is concave upward at a minimum (c) and downward at a maximum, such as a and b. As the second derivative of the function is the rate of change of the slope, the sign of the second derivative provides a method of distinguishing a minimum from a maximum. In the former case the second derivative is positive, while in the latter it is negative. The value of the second derivative at an extreme point is referred to as the curvature of the function at that point.

A case that has not yet been considered in this section is shown in Fig. 5 at the point $x = d$. At this point the slope of the first derivative is equal to zero, that is

$$\left.\frac{\mathrm{d}^2f}{\mathrm{d}x^2}\right|_{x=d} = 0.$$

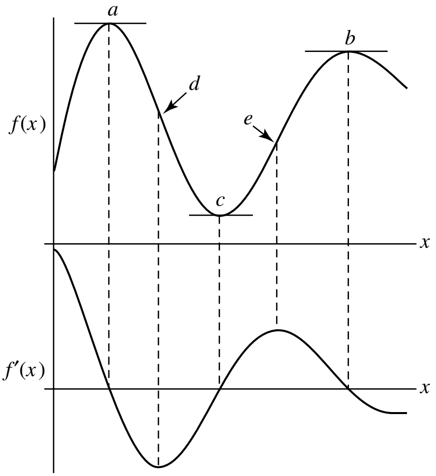


Fig. 5 A function $f(x)$ and its derivative, $f'(x)$.

Hence the point $x = d$ is neither a maximum nor a minimum of the function $f(x)$. Here, $f(x_1) > f(d) > f(x_2)$ and, as $x_1 < d < x_2$, the slope is not equal to zero. The point $x = d$ is known as an inflection point, a point at which the second derivative or curvature is zero. The point $x = e$ is also an inflection point, as $f'' = 0$. The ensemble of extrema and inflection points of a function are known as its critical points.

An example of a function which exhibits an inflection point is provided by the well-known equation of Van der Waals,* which for one mole of a gas takes the form,

$$\left(P + \frac{a}{V^2}\right)(V - b) = RT, \tag{20}$$

where a and b are constants. The derivative can be obtained in the form

$$\left(\frac{\partial P}{\partial V}\right)_T = \frac{-RT}{(V-b)^2} + \frac{2a}{V^3}, \tag{21}$$

where the subscript T indicates that the temperature has been held constant.† Note that the slope is equal to zero at infinite molar volume and becomes infinite at $V = b$. However, there is an intermediate point of interest along the curve $T = T_c$. At this, the so-called critical point, the curve exhibits an

---

*Johannes Diderik Van der Waals, Dutch physicist (1887–1923).

†The derivative in Eq. (21) is an example of a partial derivative, a subject that will be treated at the end of this chapter.

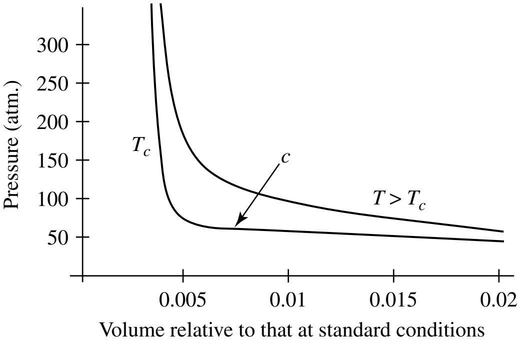


Fig. 6 Isotherms of a Van der Waals fluid; the critical point is shown at c (1 atm. = 101 kPa).

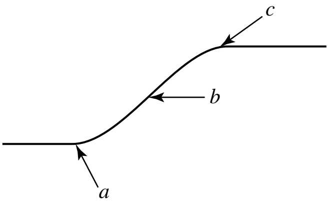


Fig. 7 Profile of a road.

inflection point, as shown in Fig. 6. At this point the derivative of Eq. (21) is equal to zero and the corresponding molar volume is given by $V_c = 3b$. The development of this result is left as an exercise (see problem 5).

It should be noted that the isotherm which passes through the critical point (Fig. 6) is a "smooth curve" in the sense that both the function $P(V)$ and its first derivative are continuous. However, the second derivative at the critical point is not.

Another, more everyday example of this behavior occurs in road construction. An automobile begins its ascent of a grade at point a in Fig. 7. The pavement is both unbroken (the function is continuous) and smooth (its derivative is continuous). However, at point a, as well as at points b and c, the second derivative, which represents the rate of change in the grade, is discontinuous.

2.7 THE DIFFERENTIAL

When the change in a variable, say Δx, approaches zero it is called an infinitesimal. The branch of mathematics known as analysis, or the calculus, is based on this principle, as both Δx and Δy approach zero in the limit [see Eq. (8)]. For practical purposes the derivative $\mathrm{d}y/\mathrm{d}x$ can be decomposed into differentials in the form $\mathrm{d}y = (\mathrm{d}y/\mathrm{d}x)\mathrm{d}x$. While this operation deserves some justification from a purely mathematical point of view, it is correct for the purposes of this book.

In this context Eq. (12) can be rewritten in the form of differentials as

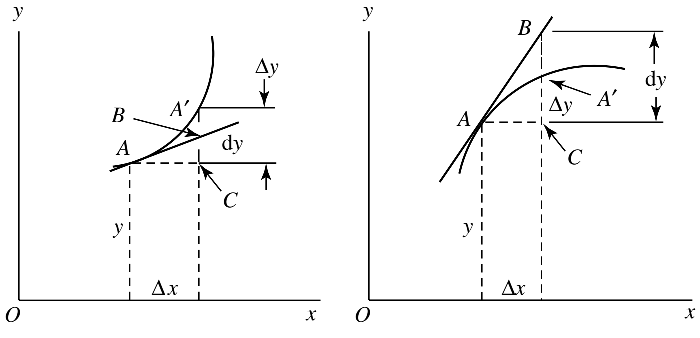
$$\mathrm{d}(uv) = u\,\mathrm{d}v + v\,\mathrm{d}u. \tag{22}$$

In other words the differential of a product of two functions is equal to the first function times the differential of the second, plus the second times the differential of the first. Numerous examples of this principle will be encountered in the exercises at the end of this chapter, as well as in following chapters. The other rules presented above can easily be modified accordingly.

A geometrical interpretation of the differential is represented in Fig. 8. It is apparent that in general $\mathrm{d}y < \Delta y$ or $\mathrm{d}y > \Delta y$, as the curve is concave upward or downward, respectively.

It is often useful to evaluate the differential along a curve s such as shown in Fig. 9. Let Δs be the length of the curve $y = f(x)$ measured between points a and b and assume that s increases as x increases. Thus, the derivative can be expressed as

$$\frac{\mathrm{d}s}{\mathrm{d}x} = \lim_{\Delta x \to 0} \frac{\Delta \sigma}{\Delta x} = \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2} \tag{23}$$



(a) (b)

Fig. 8 Geometrical interpretation of the differential.

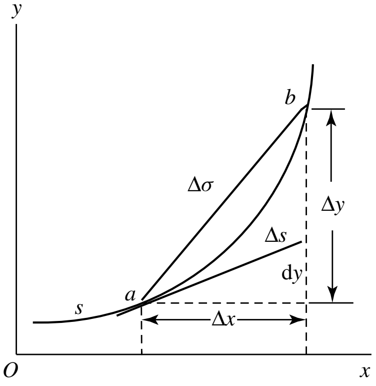


Fig. 9 Geometrical illustration of the infinitesimal of arc, d*s*.

and, as Δs approaches $\Delta\sigma$ in the limit,

$$\overline{\mathrm{d}s}^2 = \overline{\mathrm{d}x}^2 + \overline{\mathrm{d}y}^2. \tag{24}$$

Then the differential ds becomes the hypotenuse of the triangle shown in Fig. 9. The same result is obtained if s decreases as x increases.

2.8 THE MEAN-VALUE THEOREM AND L'HOSPITAL'S RULE*

An important theorem, often attributed to Lagrange,† can be written in the form

$$f'(x_1) = \frac{f(b) - f(a)}{b - a}. \tag{25}$$

Here $f'(x_1)$ is the derivative $\mathrm{d}y/\mathrm{d}x$ evaluated at a point x_1 which is intermediate with respect to points a and b. From Fig. 10 it should be evident that there is always some point $x = x_1$ where the slope of the curve is equal to the right-hand side of Eq. (25). This theorem will be employed in Chapter 13 to evaluate the error in linear interpolation.

If two functions $f(x)$ and $g(x)$ both vanish at a point a, the ratio $f(a)/g(a)$ is undefined. It is the so-called indeterminate form $\frac{0}{0}$ mentioned earlier

---

*Guillaume de L'Hospital, French mathematician (1661–1704).

†Louis de Lagrange, French mathematician (1736–1813).

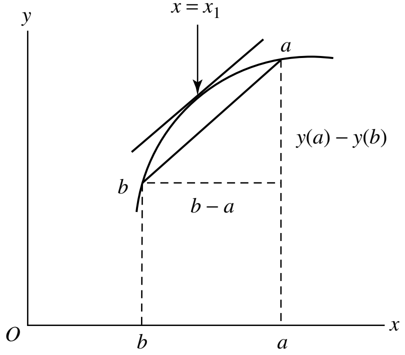


Fig. 10 The slope of a curve at $x = x_1$.

(Sections 2.1 and 2.2). However, the limit of this ratio may exist. In fact this principle is the very basis of the differential calculus, as indicated by Eq. (9).

Consider now Eq. (25), with b replaced by x, *viz.*

$$f(x) = (x - a)f'(x_1) + f(a). \tag{26}$$

Similarly, for another function $g(x)$,

$$g(x) = (x - a)g'(x_2) + g(a); \tag{27}$$

and, as the case of interest is $f(a) = g(a) = 0$, Eqs. (26) and (27) yield

$$\frac{f(x)}{g(x)} = \frac{f'(x_1)}{g'(x_2)}. \tag{28}$$

Because both x_1 and x_2 lie between x and a, they both must approach a as x does. Thus,

$$\mathit{lim}_{x \to a} \frac{f(x)}{g(x)} = \mathit{lim}_{x \to a} \frac{f'(x)}{g'(x)}, \tag{29}$$

which is known as L'Hospital's rule. A trivial example of its application is provided by the function $\mathit{sin}\, x/x$. In this case the ratio of the first derivatives evaluated at the origin is equal to unity, as shown earlier.

An example which may be familiar to chemists, as it arises in extraction and fractional distillation, is the function

$$y = \frac{x^{n+1} - x}{x^{n+1} - 1}, \tag{30}$$

where n is an integer. This function is undefined at $x = 1$. However, with the application of Eq. (29), the limit is given by

$$\lim_{x\to 1} \frac{x^{n+1}-x}{x^{n+1}-1} = \lim_{x\to 1} \frac{(n+1)x^n - 1}{(n+1)x^n} = \frac{n}{n+1} \tag{31}$$

for all finite values of n.

L'Hospital's rule has been applied above to cases in which the indeterminate form is $\frac{0}{0}$. However, it is equally valid for the form $\frac{\infty}{\infty}$.

2.9 TAYLOR'S SERIES*

Power series have already been introduced to represent a function. For example, Eq. (1-35) expresses the function $y = \sin x$ as a sum of an infinite number of terms. Clearly, for $x < 1$, terms in the series become successively smaller and the series is said to be convergent, as discussed below. The numerical evaluation of the function is carried out by simply adding terms until the value is obtained with the desired precision. All computer operations used to evaluate the various irrational functions are based on this principle.

Now assume that a given function can be differentiated indefinitely at a given point a and that its expansion in a power series is of the form

$$f(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \cdots. \tag{32}$$

If this series converges in the region around the point a, it can be used to calculate the function $f(x)$ to a precision determined by the number of terms retained. Assuming that the series exists, the coefficients can be determined. Certainly, $c_0 = f(a)$ and, by successive, term-by-term differentiation the subsequent coefficients are evaluated. Thus, as

$$f'(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + 4c_4(x-a)^3 + \cdots, \tag{33}$$

$$f'(a) = c_1,$$

and

$$\begin{gathered} f''(a) = 2\cdot 1 c_2 \\ f'''(a) = 3\cdot 2\cdot 1 c_3 \\ \vdots \\ f^{(n)}(a) = n! c_n. \end{gathered}$$

---

*Brook Taylor, British mathematician (1685–1731).

The general form of Taylor's series is then

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + \frac{1}{3!}f'''(a)(x-a)^3 + \cdots + \frac{1}{n!}f^{(n)}(a)(x-a)^n + \cdots . \tag{34}$$

Thus, it has been shown that if a series as presented in Eq. (32) exists, it is given by Eq. (34). However, the function and its successive derivatives must be defined at $x = a$. Furthermore, the function must be analytic, the series must be convergent in this region and the value obtained must be equal to $f(x)$. These questions deserve further consideration for a given problem.

An example of the development of a Taylor's series is provided by the expansion of the function *ln x* around the point $x = 1$. The necessary derivatives become

$$\begin{aligned}
f'(x) &= \frac{1}{x} & f'(1) &= 1 \\
f''(x) &= -\frac{1}{x^2} & f''(1) &= -1 \\
f'''(x) &= \frac{2}{x^3} & f'''(1) &= 2 \\
f^{(iv)}(x) &= -\frac{2\cdot 3}{x^4} & f^{(iv)}(1) &= -2\cdot 3 \\
&\vdots & &\vdots \\
f^{(n)}(x) &= (-1)^{n-1}\frac{(n-1)!}{x^n} & f^{(n)}(1) &= (-1)^{n-1}(n-1)! \\
&\vdots & &\vdots
\end{aligned}$$

and the series is then

$$ln\, x = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \cdots + (-1)^{n-1}\frac{(n-1)^n}{n} + \cdots . \tag{35}$$

It can be shown that this series converges for $0 < x \le 2$ (see Section 2.11).

An important special case of Taylor's series occurs when $a = 0$. Then, Eq. (34) takes the form

$$f(x) = f(0) + f'(0)x + \frac{1}{2!}f''(0)x^2 + \frac{1}{3!}f'''(0)x^3 + \cdots + \frac{1}{n!}f^{(n)}(0)x^n + \cdots , \tag{36}$$

which is known as Maclaurin's series.* An application was introduced in Chapter 1, where one definition of the function *sin x* was expressed as an infinite series [see Eq. (1-35) and problem 9].

2.10 BINOMIAL EXPANSION

Consider the development of the function $(x+1)^\alpha$ in a Maclaurin series,

$$f(x) = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + \cdots. \tag{37}$$

The coefficients are known in the form

$$\binom{\alpha}{n} \equiv \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} \tag{38}$$

as the binomial coefficients. In the special case in which $\alpha = n$, a positive integer,

$$\binom{n}{n} = 1$$

and

$$\binom{n}{n+1} = \binom{n}{n+2} = \cdots = 0.$$

The infinite series given by Eq. (37) then reduces to the polynomial

$$(x+1)^n = 1 + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n, \tag{39}$$

which is Newton's binomial formula.†

The binomial expansion, Eq. (37), is particularly useful in numerical applications. For example, if $\alpha = \frac{1}{2}$,

$$f(x) = (x+1)^{1/2} = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \cdots. \tag{40}$$

---

*Colin Maclaurin, Scottish mathematician (1698–1746).

†Sir Isaac Newton, English physicist and mathematician (1642–1727).

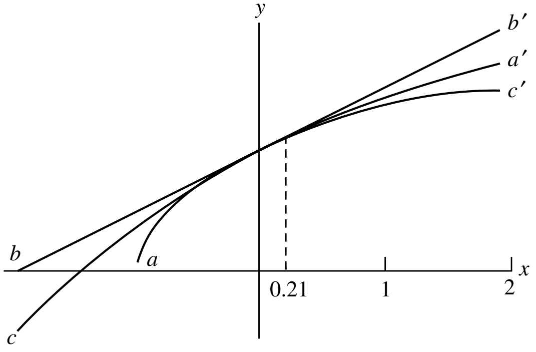


Fig. 11 (**a**) The function $y = (x+1)^{1/2}$; (**b**) Linear approximation; (**c**) Quadratic approximation.

This function is shown in Fig. 11, where it is compared with the two- and three-term approximations derived from Eq. (40). At values of x near zero these approximations become increasingly accurate. If, for example, $x = 0.2100$, $y = (1 + 0.2100)^{1/2} = 1.1000$, while the two-term approximation yields $y = 1 + 0.2100/2 = 1.1050$. This development is often employed in computer programs. Clearly, for a given value of x the number of terms used is determined by the precision required in the numerical result.

2.11 TESTS OF SERIES CONVERGENCE

The most useful test for the convergence of a series is called Cauchy's ratio test.* It can be summarized as follows for a series defined by Eq. (32).

(i) If $lim_{n\to\infty} |c_{n+1}/c_n| < 1$ the series converges absolutely,† and thus converges.

(ii) If $lim_{n\to\infty} |c_{n+1}/c_n| > 1$, or if $|c_{n+1}/c_n|$ increases indefinitely, the series diverges.

(iii) If $lim_{n\to\infty} |c_{n+1}/c_n| = 1$ or if the quantity $|c_{n+1}/c_n|$ does not approach a limit and does not increase indefinitely, the test fails.

---

*Augustin Cauchy, French mathematician (1789–1857).

†A series is said to be absolutely convergent if the series formed by replacing all of its terms by their absolute values is convergent.

As an example, consider the series

$$y = e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \tag{41}$$

which was introduced in Chapter 1 as a definition of the exponential function. Application of the ratio test yields

$$\lim_{n\to\infty} \frac{x^{n+1}/(n+1)!}{\mathrm{x}^n/n!} = \lim_{n\to\infty} \frac{x}{n} = 0\ ; \tag{42}$$

thus the series converges for all finite values of x.

Another test can be applied in the case of an alternating series, that is, one in which the terms are alternately positive and negative. It can be shown that if, after a certain number of terms, further terms do not increase in value and that the limit of the n^{th} term is zero, the series is convergent.

As an example, consider the series

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}, \tag{43}$$

which was introduced in Section 1.7 [see Eq. (1-35)]. This series is alternating, with successive terms decreasing in absolute value. Furthermore, as

$$\lim_{n\to\infty} \frac{x^{2n}}{(2n)!} = 0, \tag{44}$$

the power series which defines the sine function is convergent for all finite values of x.

Two other important considerations are involved in the use of infinite series. Convergence may be assured only within a given range of the independent variable, or even only at a single point. Thus, the “region of convergence” can be identified for a given series. The reader is referred to textbooks on advanced calculus for the analysis of this problem.

A second question arises in practical applications, because at different points within the region of convergence, the rate of convergence may be quite different. In other words the number of terms that must be retained to yield a certain level of accuracy depends on the value of the independent variable. In this case the series is not uniformly convergent.

2.12 FUNCTIONS OF SEVERAL VARIABLES

Thus far in this chapter, functions of only a single variable have been considered. However, a function may depend on several independent variables. For example, $z = f(x,y)$, where x and y are independent variables. If one of these variables, say y, is held constant, the function depends only on x. Then, the derivative can be found by application of the methods developed in this chapter. In this case the derivative is called the partial derivative of z with respect to x, which is represented by $\partial z/\partial x$ or $\partial f/\partial x$. The partial derivative with respect to y is analogous. The same principle can be applied to implicit functions of several independent variables by the method developed in Section 2.5. Clearly, the notion of partial derivatives can be extended to functions of any number of independent variables. However, it must be remembered that when differentiating with respect to a given independent variable, all others are held constant.

Higher derivatives are obtained by obvious extension of this principle. Thus,

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

as in Section 2.4 [See Eq. (16)]. It should be noted, however, that the order of differentiation is unimportant if the function $z(x,y)$ is continuous. So that

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x},$$

a relation that is important, as shown in the following chapter.

It is now of interest to define the total differential by the relation

$$\mathrm{d}z = \left(\frac{\partial z}{\partial x}\right)\mathrm{d}x + \left(\frac{\partial z}{\partial y}\right)\mathrm{d}y. \tag{45}$$

This expression is a simple generalization of the argument developed in Section 2.7. It, and its extension to functions of any number of variables, is referred to as the “chain rule”. In many applications it is customary to add one or more subscripts to the partial derivatives to specify the one or more variables that were held constant. As an example, Eq. (45) becomes

$$\mathrm{d}z = \left(\frac{\partial z}{\partial x}\right)_y \mathrm{d}x + \left(\frac{\partial z}{\partial y}\right)_x \mathrm{d}y. \tag{46}$$

This notation was suggested in Eq. (21) and is usually employed in thermodynamic applications.

2.13 EXACT DIFFERENTIALS

Equation (45) can be written in the general form

$$\delta z = M(x,y)\,\mathrm{d}x + N(x,y)\,\mathrm{d}y. \tag{47}$$

However, in the special case in which $M(x,y) = \partial z/\partial x$ and $N(x,y) = \partial z/\partial y$ the differential can be identified with that given by Eq. (45). As the order of differentiation is unimportant, the relation

$$\frac{\partial M}{\partial y} = \frac{\partial^2 z}{\partial y \partial x} = \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial N}{\partial x} \tag{48}$$

is easily obtained. The total differential, which is then said to be exact, is written dz to distinguish it from the inexact differential denoted δz. The condition for exactness, as given by Eq. (48), namely,

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \tag{49}$$

is attributed to either Cauchy or Euler, depending on the author.

In thermodynamics the eight quantities P, V, T, E, S, H, F and G are the state functions, pressure, volume, temperature, energy, entropy, enthalpy, Helmholtz* free energy and Gibbs† free energy, respectively. By definition, all of the corresponding differentials are exact (see Section 3.5). The thermodynamics of systems of constant composition can be developed with the use of any of the following sets of three state functions: E,S,V; H,S,P; F,T,V; G,T,P. Thus, for example, with $E = f(S,V)$

$$\mathrm{d}E = \left(\frac{\partial E}{\partial S}\right)_V \mathrm{d}S + \left(\frac{\partial E}{\partial V}\right)_S \mathrm{d}V. \tag{50}$$

However, the first law of thermodynamics expresses the differential dE as

$$\mathrm{d}E = \delta q + \delta w, \tag{51}$$

where the additional quantities q, the heat, and w the work have been introduced.‡ Note that these two important thermodynamic quantities are not state

---

*Hermann von Helmholtz, German physicist and physiologist (1821–1894).

†J. Willard Gibbs, American chemical physicist (1839–1903).

‡The convention adopted here is that δw is negative if work is done by the system. However, in some textbooks the first law of thermodynamics is written in the form $\mathrm{d}E = \delta q - \delta w$, in which case the work done by the system is positive.

functions; thus, their differentials are not exact. However, for a gas under reversible conditions $\delta w = -P\,\mathrm{d}V$, while from the definition of the entropy as given in Section 3.5, $\delta q = T\,\mathrm{d}S$. The resulting expression,

$$\mathrm{d}E = T\,\mathrm{d}S - P\,\mathrm{d}V \tag{52}$$

can be compared to Eq. (50) to yield the relations

$$\left(\frac{\partial E}{\partial S}\right)_V = T \text{ and } \left(\frac{\partial E}{\partial V}\right)_S = -P.$$

The application of the condition given by Eq. (49) leads to one of the four Maxwell relations,* *viz.*

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V. \tag{53}$$

The other three can be derived similarly (see problem 10 and Section 7.6).

PROBLEMS

1. Find the first derivatives of the following functions:

$y = \dfrac{x-1}{\sqrt{x^2+1}}$; Ans. $y' = \dfrac{x+1}{(x^2+1)^{3/2}}$

$y = \dfrac{1}{x^2} + \sqrt[3]{(x^3 + \cos^3 x)^2}$ Ans. $y' = -\dfrac{2}{x^3} + \dfrac{2(2x - 3\sin x\cos^2 x)}{3\sqrt[3]{x^2 + \cos^3 x}}$

$y = \dfrac{x^2+1}{4x+3}$ Ans. $y' = \dfrac{2(2x-1)(x+2)}{(4x+3)^2}$

$y = \tan(x\sin x)$ Ans. $y' = (x\cos x + \sin x)\sec^2(x\sin x)$

$y = \sec^2 x - \tan^2 x$ Ans. $y' = 0$

2. Given the curve $y = \cos x$, find the points where the tangent is parallel to the x axis. Ans. $x = k\pi$, $k = 0, \pm 1, \pm 2, \cdots$

3. Evaluate the following limits:

$\lim\limits_{\alpha\to 0} \dfrac{\sin k\alpha}{\alpha}$ Ans. k

---

*James Clark Maxwell, British physicist (1831–1879).

$\lim_{x\to 2}(x^3 - 3x)$ Ans. 2

$\lim_{x\to 2}\dfrac{x^3 - x^2 + 2x - 8}{x - 2}$ Ans. 10

$\lim_{x\to 0}\dfrac{x + \tan x}{\sin 3x}$ Ans. $\dfrac{2}{3}$

$\lim_{x\to 0}\dfrac{\sin x - x}{x^3}$ Ans. $-\dfrac{1}{6}$

4. Verify Eq. (21) and calculate the second partial derivative.

Ans. $\left(\dfrac{\partial^2 P}{\partial V^2}\right)_T = \dfrac{2RT}{(V-b)^3} - \dfrac{6a}{V^4}$

5. Show that for a Van der Waals' fluid at the critical point

$$T_c = \frac{a}{27Rb},\ V_c = 3b \text{ and } P_c = \frac{a}{27b^2}$$

6. Given the relation

$$(a-b)kt = \ln\frac{b(a-x)}{a(b-x)},$$

where $a \neq b$ are constants, find the expression for $\mathrm{d}x/\mathrm{d}t$.

Ans. $\dfrac{\mathrm{d}x}{\mathrm{d}t} = \dfrac{abk(a-b)^2 e^{k(a-b)t}}{(ae^{k(a-b)t} - b)^2}$

7. Given $x = t^2 + 2t - 4$ and $y = t^2 - t + 2$, evaluate $\mathrm{d}^2y/\mathrm{d}x^2$.

Ans. *cf.* Section 2.5.

8. If $y = A\cos kx + B\sin kx$, where A, B and k are constants, find the expression for $\mathrm{d}^2y/\mathrm{d}x^2$.

Ans. $\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2} = -k^2y$

9. Verify the series for $\cos\varphi$ and $\sin\varphi$ given by Eqs. (1-34) and (1-35), respectively.

10. Verify Eq. (53) and derive the other three Maxwell relations, namely,

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial P}{\partial T}\right)_V,\quad \left(\frac{\partial T}{\partial P}\right)_S = \left(\frac{\partial V}{\partial S}\right)_P \quad\text{and}\quad \left(\frac{\partial S}{\partial P}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_P.$$

11. Find the first partial derivatives of the function $z = 4x^2y - y^2 + 3x - 1$.

Ans. $\dfrac{\partial z}{\partial x} = 8xy + 3,\ \dfrac{\partial z}{\partial y} = 4x^2 - 2y$

12. Given $z^2 + 2zx = x^2 - y^2$, find the first partial derivatives z.

Ans. $\dfrac{\partial z}{\partial x} = \dfrac{x - z}{x + z}, \dfrac{\partial z}{\partial y} = \dfrac{-y}{x + z}$

13. Verify the development of $\mathbf{\nabla^2}$ as given in Appendix V.

14. Given the function $u = 3x^2 + 2xz - y^2$, show that $x(\partial u/\partial x) + y(\partial u/\partial y) + z(\partial u/\partial z) = 2u$.

15. If $u = ln(x^2 + y^2)$, show that $(\partial^2 u/\partial x^2) + (\partial^2 u/\partial y^2) = 0$.

16. Show that $u = e^{-at} \, cos \, bt$ is a solution to the equation $(\partial^2 u/\partial x^2) = (\partial u/\partial t)$, if the constants are chosen so that $a = b^2$.

This Page Intentionally Left Blank

3 Integration

3.1 THE INDEFINITE INTEGRAL

The derivative and the differential were introduced in Chapter 2. There, given a function, the problem was to find its derivative. In this chapter the objective is to perform the inverse operation, namely, given the derivative of a function, find the function. The function in question is the integral of the given function. It is defined by the expression*

$$f(x) = \int f'(x)\,\mathrm{d}x. \tag{1}$$

As an example, consider the function $f'(x) = x^3$. The prime on $f(x)$ indicates that this function $\mathrm{d}f(x)/\mathrm{d}x$ is the derivative of the function searched. Given the rules of differentiation [Eq. (2-15)], the function might be expected to have the form $\frac{1}{4}x^4$. This result is correct, although it should be noted that the addition of any constant to the function $\frac{1}{4}x^4$ does not change the value of the derivative, as the derivative of a constant is equal to zero. It must therefore be concluded that the indefinite integral is given by

$$\int x^3\,\mathrm{d}x = \frac{1}{4}x^4 + C. \tag{2}$$

The constant C is the constant of integration introduced in the applications presented in Chapter 1 [Eqs. (1-20) and (1-24)]. There it was indicated that the determination of this constant requires additional information, namely, the initial or boundary conditions associated with the physical problem involved. Integrals of this type are, therefore, called indefinite integrals.

---

*The notation $f(x) = \int \mathrm{d}x f'(x)$ is often employed. It is, however, ambiguous in some cases and should be avoided.

3.2 INTEGRATION FORMULAS

The general rules for obtaining an indefinite integral can be summarized as follows:

$$\text{(i)}\quad \int \mathrm{d}u = u + C, \tag{3}$$

$$\text{(ii)}\quad \int (\mathrm{d}u + \mathrm{d}v + \cdots + \mathrm{d}z) = \int \mathrm{d}u + \int \mathrm{d}v + \cdots + \int \mathrm{d}z, \tag{4}$$

$$\text{(iii)}\quad \int a\,\mathrm{d}u = a \int \mathrm{d}u \tag{5}$$

and

$$\text{(iv)}\quad \int u^n\,\mathrm{d}u = \frac{u^{n+1}}{n+1} + C,\ \text{ if } n \neq -1, \tag{6}$$

where a is a constant and n is an integer. Rule (iv) is the general power formula of integration. It is obviously the inverse of Eq. (2-15).

Some other formulas for integration can be summarized as follows:

$$\int x^{-1}\,\mathrm{d}x = ln\, x + C \tag{7}$$

$$\int e^x\,\mathrm{d}x = e^x + C \tag{8}$$

$$\int sin\, x\,\mathrm{d}x = -\,cos\, x + C \tag{9}$$

$$\int cos\, x\,\mathrm{d}x = sin\, x + C \tag{10}$$

$$\int sec^2 x\,\mathrm{d}x = tan\, x + C \tag{11}$$

$$\int sinh\, x\,\mathrm{d}x = cosh\, x + C \tag{12}$$

$$\int cosh\, x\,\mathrm{d}x = sinh\, x + C \tag{13}$$

$$\int \frac{\mathrm{d}x}{1+x^2} = tan^{-1}\, x + C \tag{14}$$

$$\int \frac{\mathrm{d}x}{\sqrt{1-x^2}} = sin^{-1}\, x + C \tag{15}$$

Many others are available in standard integral tables* and computer programs.†

3.3 METHODS OF INTEGRATION

3.3.1 Integration by substitution

Very often integrals can be evaluated by introducing a new variable. The variable of integration x is replaced by a new variable, say z, where the two are related by a well chosen formula. Thus, the explicit substitution $x = \phi(z)$ and $\mathrm{d}x = (\mathrm{d}\phi/\mathrm{d}z)\mathrm{d}z$ can be made to simplify the desired integration. As an example, consider the integral

$$\int \frac{x^3\, \mathrm{d}x}{\sqrt{x^2-a^2}}, \tag{16}$$

where a is a constant. Let $x^2 - a^2 = z^2$; and, with $\mathrm{d}x = z\, \mathrm{d}z/x$, the integral takes the form

$$\int \frac{(z^2+a^2)z\, \mathrm{d}z}{z} = \frac{1}{3}(x^2-a^2)^{3/2} + a^2(x^2-a^2)^{1/2} + C. \tag{17}$$

The expressions, $\sqrt{x^2-a^2}$, $\sqrt{x^2+a^2}$ and $\sqrt{a^2-x^2}$ occur often in the integrand. The substitution of a new independent variable for the radical should be made whenever the integrand contains a factor which is an odd integral power of x. Otherwise, the radical will reappear after the substitution.

Trigonometric substitutions are often useful in evaluating integrals. Among the many possibilities, if the integrand involves the expression $x^2 + a^2$, the substitution $x = a\, tan\, \varphi$ should be tried. Similarly, in the cases of $x^2 - a^2$ or $a^2 - x^2$, the independent variable x should be replaced by $a\, sec\, \varphi$ or $a\, sin\, \varphi$, respectively. As an example of the latter case, consider the integral

$$I = \int \frac{\mathrm{d}x}{(a^2-x^2)^{3/2}}. \tag{18}$$

---

*B. O. Peirce, *A Short Table of Integrals*, Ginn and Company, Boston, 1929.
I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products*, Academic Press, New York, 1965.

†*Mathematica*, Wolfram Research, Inc., Champaign, Ill., 1997.

The substitutions $x = a\,sin\,\varphi$ and $\mathrm{d}x = a\,cos\,\varphi\,\mathrm{d}\varphi$ yield

$$\begin{aligned} I &= \int \frac{a\,cos\,\varphi\,\mathrm{d}\varphi}{(a^2 - a^2\,sin^2\,\varphi)^{3/2}} = \frac{1}{a^2}\int \frac{cos\,\varphi\,\mathrm{d}\varphi}{cos^3\varphi} \\ &= \frac{1}{a^2}\int sec^2\varphi\,\mathrm{d}\varphi = \frac{1}{a^2}\,tan\,\varphi + C \\ &= \frac{x}{a^2\sqrt{a^2 - x^2}} + C. \end{aligned} \tag{19}$$

3.3.2 Integration by parts

This method is the direct result of Eq. (2-22) for the differential of a product,

$$\mathrm{d}(uv) = u\,\mathrm{d}v + v\,\mathrm{d}u. \tag{20}$$

Therefore,

$$\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u, \tag{21}$$

which is the basic formula for integration by parts. This method is very useful, although it is not always clear how to break up the integrand. As an example, consider the integral

$$I_1 = \int xe^{-x}\,\mathrm{d}x. \tag{22}$$

With the choice $\mathrm{d}v = e^{-x}\,\mathrm{d}x$ and $u = x$, $v = -e^{-x}$ and, of course, $\mathrm{d}u = \mathrm{d}x$. Eq. (22) then yields

$$I_1 = -xe^{-x} + \int e^{-x}\,\mathrm{d}x = -(1 + x)e^{-x} + C. \tag{23}$$

It should be apparent that in integrating $\mathrm{d}v = e^{-x}\,\mathrm{d}x$ it is not necessary to add the constant of integration, as the final result is not changed by its inclusion.

The above example can be generalized. The integral

$$I_n = \int x^n e^{-x}\,\mathrm{d}x, \tag{24}$$

where n is a positive integer, can be reduced to Eq. (23) by successive integration by parts. Thus,

$$I_n = -x^n e^{-x} + n\int x^{n-1}e^{-x}\,\mathrm{d}x = -x^n e^{-x} + nI_{n-1}, \tag{25}$$

a result which is given in all integral tables.

3.3.3 Integration of partial fractions

Another common method of integration involves partial fractions. First, it should be noted that every rational algebraic fraction can be integrated directly. A rational algebraic fraction is the ratio of two polynomials. If the polynomial in the numerator is of a lower degree than that of the denominator, or can be made so by division, the resulting fraction can be written as the sum of fractions whose numerators are constants and whose denominators are the factors of the original denominator. Fortunately, in many cases the denominator can be broken up into real linear factors, none of which is repeated. As an example, consider the integral

$$\int \frac{x+3}{x^3-x}\,\mathrm{d}x. \tag{26}$$

The integrand can be written in the form

$$\frac{x+3}{x(x+1)(x-1)} = \frac{A}{x} + \frac{B}{x+1} + \frac{D}{x-1}\,, \tag{27}$$

where A, B and D are constants. Expressing the right-hand side of Eq. (27) over a common denominator yields the relation

$$x+3 = -A + (D-B)x + (A+B+D)x^2 \tag{28}$$

and by equating coefficients of the various powers of x, $A=-3$, $B=1$ and $D=2$. The proposed integral is then given by

$$\int \frac{x+3}{x^3-x}\,\mathrm{d}x = -3\int \frac{\mathrm{d}x}{x} + \int \frac{\mathrm{d}x}{x+1} + 2\int \frac{\mathrm{d}x}{x-1} \tag{29}$$

$$= -3\,ln\,x + ln(x+1) + 2\,ln(x-1) + C$$

$$= ln\,\frac{(x+1)(x-1)^2}{x^3} + C. \tag{30}$$

Integrals involving partial fractions occur often in chemical kinetics. For example, the differential equation which represents a second-order reaction is

$$\frac{\mathrm{d}x}{\mathrm{d}t} = k(a-x)(b-x), \tag{31}$$

where k is the rate constant and a and b are the initial concentrations of the two reactants. In Eq. (31) the independent variable x represents the concentration of product formed at time t. After separation of variables, Eq. (31) becomes

$$\frac{\mathrm{d}x}{(a-x)(b-x)} = k\,\mathrm{d}t. \tag{32}$$

In the general case in which $a \neq b$ the integration of the left-hand side of Eq. (32) can be carried out with the use of partial fractions. Then, the integrand is broken up in the form

$$\frac{1}{(a-x)(b-x)} = \frac{A}{(a-x)} + \frac{B}{(b-x)} \ , \tag{33}$$

with $A = 1/(b-a)$ and $B = -1/(b-a)$. The integral of Eq. (31) can then be expressed as

$$\int \frac{\mathrm{d}x}{(a-x)(b-x)} = \frac{1}{(b-a)}[-\,ln(a-x) + ln(b-x)] + C = kt. \tag{34}$$

The initial condition $x = 0$ at $t = 0$ leads to the value of the integration constant, *viz.*

$$C = \frac{1}{(b-a)}(ln\, b - ln\, a) = \frac{1}{(a-b)}\, ln\, \frac{a}{b} \ , \tag{35}$$

and the resulting expression for the concentration of product at time t,

$$x = \frac{ab(1 - e^{kt(a-b)})}{b - ae^{kt(a-b)}}. \tag{36}$$

The integration method illustrated above becomes somewhat more complicated if the denominator contains repeated linear factors. Thus, if the denominator contains a factor such as $(x-a)^n$, n identical factors would result which could of course be combined. To avoid this problem it is assumed that $1/(x-a)^n$ can be replaced by

$$\frac{A}{x-a} + \frac{B}{(x-a)^2} + \cdots + \frac{N}{(x-a)^n}. \tag{37}$$

The constants appearing in the numerator are then evaluated as before.

The differential equation for a chemical reaction of third order is of the general form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = k(a-x)(b-x)(c-x), \tag{38}$$

where a, b and c are the initial concentrations of the three reactants. It can be integrated directly by application of the method just illustrated. In the special case in which two of the reactants have the same initial concentration, say $b = c$, Eq. (38) becomes

$$\frac{\mathrm{d}x}{\mathrm{d}t} = k(a-x)(b-x)^2 \tag{39}$$

and the integral to be evaluated is

$$\int \frac{\mathrm{d}x}{(a-x)(b-x)^2} \; . \tag{40}$$

As the factor $b - x$ appears twice in the denominator, the partial fractions must be developed as given by Eq. (37), namely

$$\frac{1}{(a-x)(b-x)^2} = \frac{A}{a-x} + \frac{B}{b-x} + \frac{D}{(b-x)^2} \; . \tag{41}$$

The constants in the numerator are found to be $A = 1/(a-b)^2$, $B = -1/(a-b)^2$ and $D = 1/(a-b)$, yielding

$$\int \frac{\mathrm{d}x}{(a-x)(b-x)^2} = \frac{1}{(a-b)^2}\, ln\, \frac{b-x}{a-x} + \frac{1}{(a-b)(b-x)} = kt + C. \tag{42}$$

The evaluation of the constant of integration is achieved by applying the initial condition $x = 0$ at $t = 0$.

3.4 DEFINITE INTEGRALS

3.4.1 Definition

Let $f(x)$ be a function whose integral is $F(x)$ and a and b two values of x. The change in the integral, $F(b) - F(a)$, is called the definite integral of $f(x)$ between the limits a and b. It is represented by

$$\int_a^b f(x)\,\mathrm{d}x = [F(x)]_a^b = F(b) - F(a) \tag{43}$$

and it is evident that the constant of integration cancels.

All definite integrals have the following two properties:

$$\text{(i)} \int_a^b f(x)\,\mathrm{d}x = -\int_b^a f(x)\,\mathrm{d}x \tag{44}$$

and

$$\text{(ii)} \int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x. \tag{45}$$

Relation (ii) is useful in the case of a discontinuity, *e.g.* a missing point at c, which usually lies between a and b.

3.4.2 Plane area

For simplicity, assume that a continuous function $f(x)$ is divided into n equal intervals of width Δx (see Fig. 1). Each rectangle of width Δx at a given point $f(x)$ has an area of $f(x)\Delta x$. Therefore, the definition of the area $\mathcal{A}$ bounded by the curve $y = f(x)$, the x axis and the limits $x = a$ and $x = b$ is given by

$$\mathcal{A} = \lim_{\Delta x \to 0} \sum_{k=1}^{n} f(x_k)\Delta x_k \ . \tag{46}$$

Thus, if Δx_k is taken to be sufficiently small, and the number of rectangles correspondingly large, the sum of the areas of the rectangles will approximate, to the desired degree of accuracy, the value of the area $\mathcal{A}$. Thus, as the widths Δx_k approach zero, the number of them, n, must approach infinity. It should be noted here that the intervals Δx_k have been assumed to be constant over the range a,b. It is not necessary from a fundamental point of view to divide the abscissa in equal steps Δx_k, although in most numerical calculations it is essential, as shown in Chapter 13.

Assuming that the required limit exists and that it can be calculated, the fundamental theorem of the integral calculus can be stated as follows.

$$\lim_{\Delta x \to 0} \sum_{k=1}^{n} f(x_k)\Delta x_k = \int_a^b f(x)\,\mathrm{d}x \tag{47}$$

and the desired area is given by

$$\mathcal{A} = \int_a^b f(x)\,\mathrm{d}x. \tag{48}$$

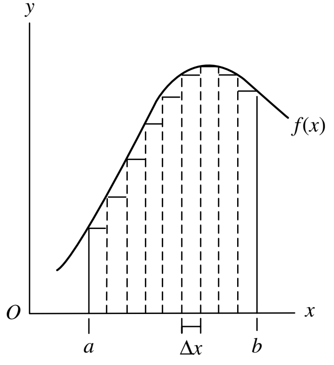


Fig. 1 The integral from $x = a$ to $x = b$.

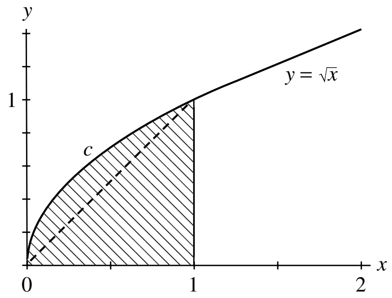


Fig. 2 The area under the curve c [Eq. (48)] and the length of the curve c [Eq. (54)].

As an example, consider the parabola $y = \sqrt{x}$ shown in Fig. 2. The area in the first quadrant under the curve between $x = 0$ and $x = 1$ is equal to

$$\int_0^1 x^{1/2}\,\mathrm{d}x = \left[\frac{2}{3}x^{3/2}\right]_0^1 = \frac{2}{3}, \tag{49}$$

as shown by the shaded area in Fig. 2. If the curve in this figure were cut into equal horizontal "slices" of width dy, the same area could be calculated as

$$\int_0^1 \mathrm{d}y - \int_0^1 y^2\,\mathrm{d}y = 1 - \frac{1}{3} = \frac{2}{3}, \tag{50}$$

where the first term corresponds to the square of unit area.

3.4.3 Line integrals

In Section 2.7 it was shown [see Eq. (2-23)] that a given element ds along a curve is given by

$$\mathrm{d}s = \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x. \tag{51}$$

Thus,

$$s = \int_a^b {}_c\,\mathrm{d}s = \int_a^b \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x. \tag{52}$$

The symbol $\int_a^b {}_c\,\mathrm{d}s$ indicates that the integral is taken along the curve c from the point a to the point b. If the variables x and y are related *via* a parameter t, the length of the curve can also be evaluated from the equivalent relation

$$s = \int_{t_a}^{t_b} \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t \tag{53}$$

where t_a and t_b are the values of t at points a and b, respectively.

As an example, consider the passage from the origin to the point (1,1), as shown in Fig. 2. Obviously, the length of a straight line between these two points, the dotted line, is equal to $\sqrt{2} = 1.414$. However, from Eq. (52) the length of the curve defined by the parabola $y = \sqrt{x}$ between these same two points is given by

$$\int_0^1 \sqrt{1 + \frac{1}{4x}}\,\mathrm{d}x = \frac{1}{4}\left[z\sqrt{1+z^2} + ln\left(z + \sqrt{1+z^2}\right)\right]_0^2 = 1.478, \tag{54}$$

where the substitution $z = 2\sqrt{x}$ has been made to simplify the integration. It should be noted that the upper limit to the integral is at $x = 1$, where $z = 2$.

The method illustrated here for determining the length of a given curve can be extended to evaluate the surface of a solid. It is particularly useful in engineering applications to determine, for example, the surface generated by the revolution of a given contour.

3.4.4 Fido and his master

To illustrate some of the principles outlined above, consider the following story. A jogger leaves a point taken as the origin in Fig. 3 at a constant speed equal to v. His dog, Fido, is at that moment at the point $x = a$. As the jogger continues in the y direction, Fido runs twice as fast, at a speed $2v$, always headed towards his master. The problem is then to find the equation that represents Fido's trajectory and the time at which he meets his master. The answers to these two questions are indicated in Fig. 3. The solution is as follows.

The distance along the y axis covered by the jogger at time t is $y = vt$. Thus, the slope of the curve followed by Fido is given by

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{vt - y}{x}. \tag{55}$$

At the same time Fido has traveled a distance $s = 2vt$ along the curve. Therefore, replacing vt by $s/2$ in Eq. (55) and rearranging, yields

$$x\frac{\mathrm{d}y}{\mathrm{d}x} = y - \frac{s}{2}\ . \tag{56}$$

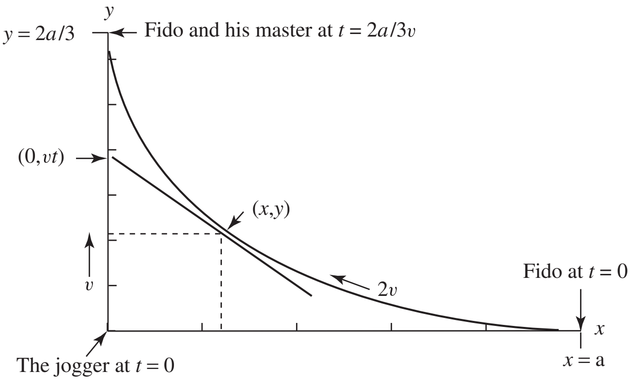


Fig. 3 Fido and his master.

With the use of Eq. (51) and the derivative of Eq. (56),

$$\frac{\mathrm{d}s}{\mathrm{d}x} = -2x\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}, \tag{57}$$

where the negative sign is chosen on the radical because the distance covered by Fido increases as x decreases. Equation (57) can be rewritten as a first order differential equation for $y' \equiv \mathrm{d}y/\mathrm{d}x$. The variables are separable, *viz.*

$$\frac{\mathrm{d}y'}{\sqrt{1+y'^2}} = \frac{\mathrm{d}x}{2x}. \tag{58}$$

The integration of the left-hand side of Eq. (58) can be carried out with the aid of the substitution $y' = \tan\theta$ (as suggested in Section 3.3.1), and the tabulated integral $\int \sec\theta\,\mathrm{d}\theta = \ln(\tan\theta + \sec\theta) + C$. The result is

$$\ln\left(y' + \sqrt{1+y'^2}\right) + C = \tfrac{1}{2}\ln x. \tag{59}$$

If the initial condition $y' = 0$ at $t = 0$ (where $x = a$) is imposed, the constant of integration is $C = \frac{1}{2}\ln a$, and the solution becomes

$$\sqrt{1+y'^2} = \sqrt{\frac{x}{a}} - y'. \tag{60}$$

Here, again, the variables can be separated. If the left-hand side of Eq. (60) is multiplied and divided by $y' - \sqrt{1+y'^2}$, the relation

$$y' - \sqrt{1+y'^2} = -\sqrt{\frac{a}{x}} \tag{61}$$

is easily obtained. Elimination of the radical $\sqrt{1+y'^2}$ between Eqs. (60) and (61) leads to

$$y' \equiv \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{2}\left(\sqrt{\frac{x}{a}} - \sqrt{\frac{a}{x}}\right), \tag{62}$$

which integrates to

$$y = \frac{x^{3/2}}{3\sqrt{a}} - \sqrt{ax} + \frac{2a}{3}. \tag{63}$$

The boundary condition $x = a$ at $y = 0$ has been employed to evaluate the second integration constant. It is Eq. (63) that describes Fido's path, as shown in Fig. 3. At $x = 0$ he "catches up" with his master, who has jogged a distance equal to $2a/3$ in time $2a/3v$.

3.4.5 The Gaussian and its moments

A very important example of integration by substitution, is that of the function of Gauss,* $exp(-z^2)$, which is shown in Fig. 4. In practical applications this function can be written in the form

$$f(z) = \mathcal{N} e^{-\alpha z^2}, \tag{64}$$

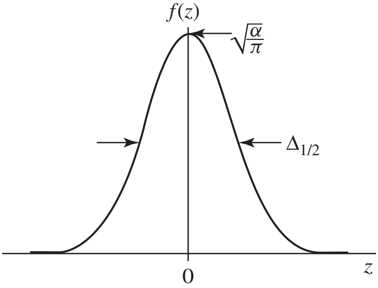


Fig. 4 The Gaussian function.

---

*Carl Friedrich Gauss, German astronomer and mathematician (1777–1855).

where α is a constant and the factor $\mathcal{N}$ is chosen to normalize the function. The latter can be evaluated by application of the normalization condition

$$\int_{-\infty}^{+\infty} f(z)\,\mathrm{d}z = 1. \tag{65}$$

Then,*

$$\mathcal{N} \int_{-\infty}^{+\infty} e^{-\alpha z^2}\,\mathrm{d}z = \mathcal{N} \cdot 2I_0 = 1 \tag{66}$$

with the definition of the integrals

$$I_n \equiv \int_0^{\infty} z^n e^{-\alpha z^2}\,\mathrm{d}z \tag{67}$$

for $n = 0, 1, 2, \ldots$. Clearly, the integral I_0 has the same value for any choice of symbol for the independent variable, say x or y. Thus,

$$I_0^2 = \int_0^{+\infty} e^{-\alpha x^2}\,\mathrm{d}x \cdot \int_0^{+\infty} e^{-\alpha y^2}\,\mathrm{d}y = \int_0^{+\infty}\int_0^{+\infty} e^{-\alpha(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y. \tag{68}$$

Equation (68) can be converted to polar coordinates with the substitutions $x = r\,cos\,\theta$ and $y = r\,sin\,\theta$. The result is given by

$$I_0^2 = \int_0^{+\infty}\int_0^{\pi/2} e^{-\alpha r^2} r\,\mathrm{d}\theta\,\mathrm{d}r = \frac{\pi}{4\alpha} \tag{69}$$

and from Eq. (66) $\mathcal{N} = \sqrt{\alpha/\pi}$. The normalized Gaussian function of Eq. (64) is then

$$f(z) = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha z^2}. \tag{70}$$

In certain applications it is of interest to express the width of the Gaussian distribution at half its value at the maximum. Thus, as the maximum value is $f(0) = \sqrt{\alpha/\pi}$, the value of z at half of this value is $\sqrt{ln\,2/\alpha}$ and the width at half-height is given by

$$\Delta_{1/2} = 2\sqrt{\frac{ln\,2}{\alpha}}, \tag{71}$$

as indicated in Fig. 4. The quantity $\Delta_{1/2}$ is referred to by spectroscopists as the "FWHM" for full width at half-maximum.

---

*Note that $f(x)$ is an even function (Section 1.2); $\int_{-\infty}^{+\infty} f(z)\,\mathrm{d}z = 2\int_0^{+\infty} f(z)\,\mathrm{d}z$.

A more general distribution can be expressed by a series of so-called moments, which are defined by

$$M(n) \equiv \int_{-\infty}^{+\infty} z^n f(z)\,\mathrm{d}z. \tag{72}$$

In the present example $f(z)$ is an even function, as given by Eq. (70). Hence, all moments with odd values of n vanish and the distribution is symmetric with respect to the origin. The first moment in this case is given by

$$\begin{aligned} M(1) &= \int_{-\infty}^{+\infty} zf(z)\,\mathrm{d}z = \int_{-\infty}^{0} zf(z)\,\mathrm{d}z + \int_{0}^{+\infty} zf(z)\,\mathrm{d}z \\ &= -\int_{0}^{-\infty} zf(z)\,\mathrm{d}z + \int_{0}^{+\infty} zf(z)\,\mathrm{d}z = 0. \end{aligned} \tag{73}$$

The second moment is

$$M(2) \equiv \int_{-\infty}^{+\infty} z^2 f(z)\,\mathrm{d}z = 2\int_{0}^{+\infty} z^2 f(z)\,\mathrm{d}z, \tag{74}$$

which, with $f(z)$ given by Eq. (70), becomes

$$M(2) = 2\int_{0}^{+\infty} z^2 e^{-\alpha z^2}\,\mathrm{d}z = \frac{1}{2\alpha}\sqrt{\frac{\pi}{\alpha}}. \tag{75}$$

The integral in Eq. (75) can be evaluated by parts (see problem 4).

3.5 INTEGRATING FACTORS

The concept of the total differential was introduced in Section 2.12. It is of importance in many physical problems and in particular in thermodynamics. In this application it is often necessary to integrate an expression of the form

$$\delta z = M(x, y)\,\mathrm{d}x + N(x, y)\,\mathrm{d}y \tag{76}$$

to determine the function $z(x,y)$ evaluated at two points, x_1,y_1 and x_2,y_2. In general this integration requires the knowledge of a relation between x and y, *i.e.* $y = f(x)$. Such a function specifies the path between the two points, and the integral becomes a line integral. As shown in Section 3.4.3, the value of the integral then depends on the chosen path.*

---

*If a differential such as given in Eq. (76) is not exact, it is represented by δz, following the custom in thermodynamics [see Eq. (2-47)].

If the differential dz is exact, according to the chain rule it is given by Eq. (2-45), *viz.*

$$\mathrm{d}z = \left(\frac{\partial z}{\partial x}\right) \mathrm{d}x + \left(\frac{\partial z}{\partial y}\right) \mathrm{d}y, \tag{77}$$

with

$$M(x, y) = \frac{\partial z}{\partial x} \text{ and } N(x, y) = \frac{\partial z}{\partial y}. \tag{78}$$

It is evident that $z(x,y)$ can be found even if the functional relation between x and y is unknown. In this case, then, the integral is independent of the path, as it depends only on the values of x and y at the two limits.

In thermodynamic applications the integral is often taken around a closed path. That is, the initial and final points in the x,y plane are identical. In this case the integral is equal to zero if the differential involved is exact, and different from zero if it is not. In mechanics the former condition defines what is called a conservative system (see Section 4.14).

Equation (76) can be written as

$$\mathrm{d}z = \mu(x, y)\delta z = \mu(x, y)M(x, y)\,\mathrm{d}x + \mu(x, y)N(x, y)\,\mathrm{d}y, \tag{79}$$

where $\mu(x, y)$ is an integrating factor. It should be noted that the integrating factor is not unique, as there is an infinite choice. In general, it is sufficient to find one suitable factor for the problem at hand.

An example is provided by the differential $y\,\mathrm{d}x - x\,\mathrm{d}y$, which is not exact. It is therefore written in the form

$$\delta z = y\,\mathrm{d}x - x\,\mathrm{d}y. \tag{80}$$

However, if this expression is multiplied by $1/y^2$, it becomes

$$\frac{y\,\mathrm{d}x - x\,\mathrm{d}y}{y^2} = \mathrm{d}\left(\frac{x}{y}\right), \tag{81}$$

which is an exact differential. Clearly, $1/y^2$ can be identified as an integrating factor.

In thermodynamics the first law is often written in the form (see Section 2.13)

$$\mathrm{d}E = \delta q + \delta w, \tag{82}$$

where dE is the (exact) differential of the internal energy of a system, while δq and δw are the (inexact) differentials of the heat and work, respectively. To illustrate the roles of exact and inexact differentials consider the work done by the reversible expansion or compression of a gas, as given by the expression

$$\delta w = -P\,\mathrm{d}V. \tag{83}$$

Equation (82) then yields

$$\delta q = \mathrm{d}E + P\mathrm{d}V. \tag{84}$$

However, as $E = f(V,T)$,

$$\mathrm{d}E = \left(\frac{\partial E}{\partial V}\right)_T \mathrm{d}V + \left(\frac{\partial E}{\partial T}\right)_V \mathrm{d}T \tag{85}$$

and

$$\delta q = \left(\frac{\partial E}{\partial T}\right)_V \mathrm{d}T + \left[P + \left(\frac{\partial E}{\partial V}\right)_T\right]\mathrm{d}V. \tag{86}$$

In this form Eq. (86) cannot be integrated without a relation between P and V, because the second term on the right-hand side involves both variables. However, in the special case in which the gas is ideal, $PV = RT$ for one mole and $(\partial E/\partial V)_T = 0$ (see problem 6). The latter relation implies the absence of intermolecular forces. Then, Eq. (86) becomes

$$\delta q = \left(\frac{\partial E}{\partial T}\right)_V \mathrm{d}T + \frac{RT\,\mathrm{d}V}{V} = \tilde{C}_V\,\mathrm{d}T + \frac{RT\,\mathrm{d}V}{V}, \tag{87}$$

where the definition of the heat capacity per mole at constant volume, $\tilde{C}_V = (\partial E/\partial T)_V$ has been introduced. While Eq. (87) can be integrated if the temperature is held constant, a more general relation is obtained by dividing by T. Thus, Eq. (87) becomes

$$\frac{\delta q}{T} = \frac{\tilde{C}_V}{T}\mathrm{d}T + \frac{R\,\mathrm{d}V}{V}. \tag{88}$$

Clearly, the differential obtained, namely, $\mathrm{d}S \equiv \delta q/T$ is exact and S, the entropy, is a thermodynamic state function, that is, it is independent of the path of integration. While Eq. (88) was obtained with the assumption of an ideal gas, the result is general if reversible conditions are applied.

With the definition of the entropy, the substitution $\delta q = T\,\mathrm{d}S$ can be made in Eq. (84); then,

$$\mathrm{d}E = T\,\mathrm{d}S - P\,\mathrm{d}V. \tag{89}$$

This result leads to one of Maxwell's relations [Eq. (2-53)]. The three remaining relations are found by analogous derivations.

3.6 TABLES OF INTEGRALS

Many tables of indefinite and definite integrals have been published. They range from collections of certain common integrals presented in appendices to most elementary calculus books, the famous Peirce tables, to compendia such as that by Gradshteyn and Ryzhik. More recently, many integrals have become available in analytical form in computer programs. One of the most complete lists is included in *Mathematica* (see footnote in Section 3.2).

Consider, as an example, the calculation of the mean-square speed of an ensemble of molecules which obey the Maxwell–Boltzmann distribution law.* This quantity is given by

$$\overline{u^2} = 4\pi(m/2\pi kT)^{3/2}\int_0^\infty e^{-mu^2/2kT}u^4\,\mathrm{d}u, \tag{91}$$

where u is the speed of a molecule of mass m, k is the Boltzmann constant and T the absolute temperature. While this integral can be evaluated by successive integration by parts (Section 3.3.2), it is much easier to employ the standard integral. It is given in the tables in the form†

$$\int_0^\infty x^{2n}e^{-ax^2}\,\mathrm{d}x = \frac{(2n-1)!!}{2(2a)^n}\sqrt{\frac{\pi}{a}}, \tag{92}$$

where n is a positive integer and $a > 0$. Comparison of Eqs. (91) and (92) allows the identifications $x = u$ and $a = m/2kT$ to be made. With $n = 2$, the integral in Eq. (91) becomes $3(\pi/2)^{1/2}(kT/m)^{5/2}$, leading to $\overline{u^2} = 3kT/m$. It should be pointed out, however, that most relatively simple integrals which can be evaluated by the methods outlined in Section 3.3 are not included in the standard tables.

When all else fails, recourse to numerical methods is indicated. Some of the classic methods of numerical integration are described in Chapter 13. However, it should be emphasized that numerical methods are to be used as a last resort. Not only are they subject to errors (often not easily evaluated), but they do not yield analytical results that can be employed in further derivations (see p. 45).

---

*Ludwig Boltzmann, Austrian physicist (1844–1906).

†The notation $(2n-1)!! = 1\cdot 3\cdot 5\cdot 7\cdots$ is often employed in integration tables.

PROBLEMS

1. Evaluate the following indefinite integrals:

$$\int \frac{(\sqrt{x}-1)^2}{\sqrt{x}}\,\mathrm{d}x \qquad \text{Ans. } \frac{2}{3}x^{3/2} - 2x + 2\sqrt{x} + C$$

$$\int \sin^3 x\,\mathrm{d}x \qquad \text{Ans. } \frac{1}{3}\cos^3 x - \cos x + C$$

$$\int \frac{\sqrt{x}\,\mathrm{d}x}{1+x} \qquad \text{Ans. } 2\sqrt{x} - 2\tan^{-1}\sqrt{x} + C$$

$$\int \frac{x^3+2}{x^3-x}\,\mathrm{d}x \qquad \text{Ans. } x - 2\ln x + \frac{1}{2}\ln(x+1) + \frac{3}{2}\ln(x-1) + C$$

2. Verify Eq. (54).

3. Verify all of the steps in the solution to the problem of Fido and his master.

4. Calculate the second moment of a Gaussian function as given by Eq. (75).

5. Show that y^{-2} is an integrating factor for the differential given by Eq. (82).

6. Demonstrate that $(\partial E/\partial V)_T = 0$ for an ideal gas.

7. Evaluate the following definite integrals:

$$\int_0^{\pi/3} x\sin x\,\mathrm{d}x \qquad \text{Ans. } 0.342$$

$$\int_{-\infty}^{\infty} \frac{x\,\mathrm{d}x}{(x^2+1)^2} \qquad \text{Ans. } 0$$

$$\int_0^{\infty} \frac{x\,\mathrm{d}x}{(x^2+1)^2} \qquad \text{Ans. } \frac{1}{2}$$

$$\int_{-1}^{+1} \frac{\mathrm{d}x}{\sqrt{4-x^2}} \qquad \text{Ans. } \pi/3$$

$$\int_0^{1} xe^{-x^2}\,\mathrm{d}x \qquad \text{Ans. } \frac{1}{2}(1-e^{-1})$$

8. Verify Eq. (92) for $n = 1$.

9. Show that the FWHM of a Gaussian is given by Eq. (71).

10. Derive an expression for I_1/I_0, where $I_0 = \int_0^\pi e^{-a\cos\theta} \sin\theta \, \mathrm{d}\theta$ and $I_1 = \int_0^\pi \cos\theta e^{-a\cos\theta} \sin\theta \, \mathrm{d}\theta$. Ans.* $\mathcal{L}(a) \equiv \coth a - \frac{1}{a}$

11. Show that $\mathcal{L}(a) \approx \frac{a}{3}$ if $a \ll 1$.†

---

*The function $\mathcal{L}(a)$ is known as the Langevin function, after Paul Langevin, French physicist (1872–1946). The magnetic susceptibility of a paramagnetic substance can be expressed as $\mathcal{L}(\mu_m \mathscr{B}/kT)$, where μ_m is the magnetic moment, $\mathscr{B}$ the magnetic flux, k the Boltzmann constant and T the absolute temperature.

†At ordinary temperatures the magnetic susceptibility is given approximately by $\mu_m \mathscr{B}/3kT$. This relation was determined experimentally by Pierre Curie, French physicist (1859–1906).

4 Vector Analysis

4.1 INTRODUCTION

To provide a mathematical description of a particle in space it is essential to specify not only its mass, but also its position (perhaps with respect to an arbitrary origin), as well as its velocity (and hence its momentum). Its mass is constant and thus independent of its position and velocity, at least in the absence of relativistic effects. It is also independent of the system of coordinates used to locate it in space. Its position and velocity, on the other hand, which have direction as well as magnitude, are vector quantities. Their descriptions depend on the choice of coordinate system. In this chapter Heaviside's notation will be followed,[*] *viz.* a scalar quantity is represented by a symbol in plain italics, while a vector is printed in bold-face italic type.

A useful image of a vector, which is independent of the notion of a coordinate system, is simply an arrow in space. The length of the arrow represents the magnitude of the vector, while its orientation in space specifies the direction of the vector. By convention the tail of the arrow is the origin of the (positive) vector and the head its terminus.

Although it is not at all necessary to describe a vector with reference to a system of coordinates, it is often useful to do so. The vector $\boldsymbol{A}$ shown in Fig. 1 represents the same quantity in either case. However, when attached to an origin (or any other given point) it can be expressed in terms of its components, which are its projections along a given set of coordinate axes. In the case of a Cartesian system[†] the magnitude of the vector $\boldsymbol{A}$, the length of the arrow, is given by

$$A = \sqrt{A_x^2 + A_y^2 + A_z^2}. \qquad (1)$$

---

[*] Oliver Heaviside, British mathematician (1850–1925).

[†] René Descartes, French philosopher, mathematician (1596–1650).

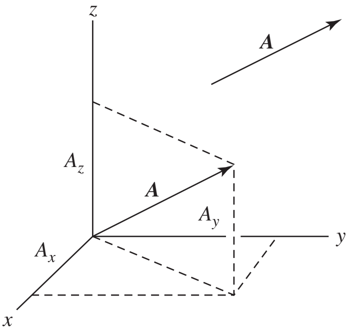


Fig. 1 A vector $\boldsymbol{A}$ in space and in a Cartesian coordinate system.

4.2 VECTOR ADDITION

The basic algebra of vectors is formulated with the aid of geometrical arguments. Thus, the sum of two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$, can be obtained as shown in Fig. 2. To add $\boldsymbol{B}$ to $\boldsymbol{A}$, the origin of $\boldsymbol{B}$ is placed at the head of $\boldsymbol{A}$ and the vector sum, represented by $\boldsymbol{R}$, is constructed from the tail of $\boldsymbol{A}$ to the head of $\boldsymbol{B}$. Clearly, the addition of $\boldsymbol{A}$ to $\boldsymbol{B}$ yields the same result (see Fig. 2); hence,

$$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A} = \boldsymbol{R} \tag{2}$$

and vector addition is commutative.

When three vectors $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are added, the resultant $\boldsymbol{R}$ is the diagonal of the parallelepiped whose edges are the vectors, as shown in Fig. 3. The same result is obtained if any two of the vectors are combined and the sum is added to the third. Thus,

$$(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{C} + \boldsymbol{A}) + \boldsymbol{B} = \boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C} = \boldsymbol{R}, \tag{3}$$

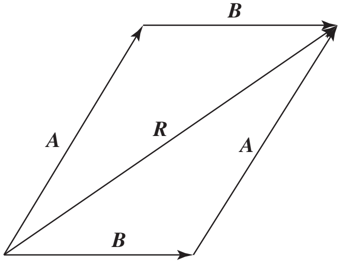


Fig. 2 The vector sum $\boldsymbol{R} = \boldsymbol{A} + \boldsymbol{B}$.

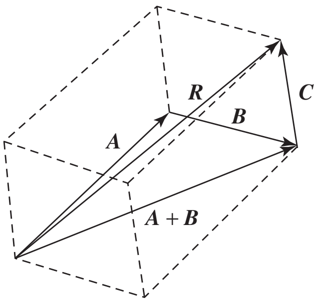


Fig. 3 The vector sum $\boldsymbol{R} = \boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C}$.

and the associative law holds. Obviously, to subtract $\boldsymbol{B}$ from $\boldsymbol{A}$, minus $\boldsymbol{B}$ is added to $\boldsymbol{A}$, *viz.*

$$\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A} + (-\boldsymbol{B}). \tag{4}$$

It should be noted that in the above presentation of the combination of vectors by addition or subtraction, no reference has been made to their components, although this concept was introduced in the beginning of this chapter. It is, however, particularly useful in the definition of the product of vectors and can be further developed with the use of unit vectors. In the Cartesian system employed in Fig. 1 the unit vectors can be defined as shown in Fig. 4. It is apparent that

$$\boldsymbol{A} = A_x\boldsymbol{i} + A_y\boldsymbol{j} + A_z\boldsymbol{k} \tag{5}$$

and similarly for another vector

$$\boldsymbol{B} = B_x\boldsymbol{i} + B_y\boldsymbol{j} + B_z\boldsymbol{k}. \tag{6}$$

The sum of these vectors is then given by

$$A_x\boldsymbol{i} + B_x\boldsymbol{i} = (A_x + B_x)\boldsymbol{i} \tag{7}$$

etc. for the other components. Then,

$$\boldsymbol{A} + \boldsymbol{B} = (A_x + B_x)\boldsymbol{i} + (A_y + B_y)\boldsymbol{j} + (A_z + B_z)\boldsymbol{k}, \tag{8}$$

where the magnitudes of parallel vectors have been added as scalars. In other words the components of the vectors can be added to obtain the components of their sums.

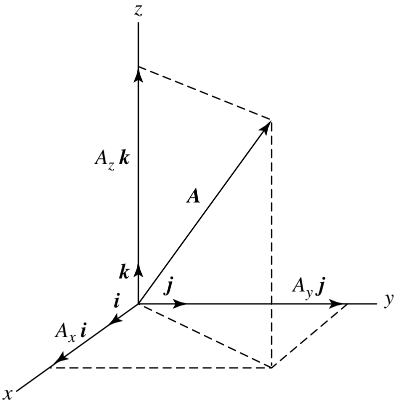


Fig. 4 Definition of the unit vectors $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$.

4.3 SCALAR PRODUCT

The scalar (or inner) product of two vectors is defined by the relation

$$\boldsymbol{A}\cdot\boldsymbol{B} = AB\,cos\,\theta, \tag{9}$$

where θ is the angle between the vectors $\boldsymbol{A}$ and $\boldsymbol{B}$. Therefore, the scalar product of two perpendicular vectors must vanish, as $\theta = \pi/2$ and $cos\,\theta = 0$. Similarly, the scalar product of any unit vector with itself must be equal to unity, as $\theta = 0$ and, hence, $cos\,\theta = 1$. In terms of the unit vectors shown in Fig. 4,

$$\boldsymbol{i}\cdot\boldsymbol{j} = \boldsymbol{j}\cdot\boldsymbol{k} = \boldsymbol{k}\cdot\boldsymbol{i} = 0 \tag{10}$$

and

$$\boldsymbol{i}\cdot\boldsymbol{i} = \boldsymbol{j}\cdot\boldsymbol{j} = \boldsymbol{k}\cdot\boldsymbol{k} = i^2 = j^2 = k^2 = 1. \tag{11}$$

From Eqs. (9) and (11) it is evident that

$$\boldsymbol{A}\cdot\boldsymbol{B} = A_xB_x + A_yB_y + A_zB_z \tag{12}$$

in a Cartesian coordinate system.

The scalar product, often called the “dot product”, obeys the commutative and distributive laws of ordinary multiplication, *viz.*

$$\boldsymbol{A}\cdot\boldsymbol{B} = \boldsymbol{B}\cdot\boldsymbol{A} \tag{13}$$

and

$$\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{C}) = (\boldsymbol{A}\cdot\boldsymbol{B}) + (\boldsymbol{A}\cdot\boldsymbol{C}). \tag{14}$$

Furthermore, it is seen from Eq. (9) that any relation involving the cosine of an included angle may be written in terms of the scalar product of the vectors which define it. Finally, the reader is warned that a relation such as $\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{A} \cdot \boldsymbol{C}$ does not imply that $\boldsymbol{B} = \boldsymbol{C}$, as

$$\boldsymbol{A} \cdot \boldsymbol{B} - \boldsymbol{A} \cdot \boldsymbol{C} = \boldsymbol{A} \cdot (\boldsymbol{B} - \boldsymbol{C}) = 0. \tag{15}$$

Thus, the correct conclusion is that $\boldsymbol{A}$ is perpendicular to the vector $\boldsymbol{B} - \boldsymbol{C}$.

4.4 VECTOR PRODUCT

Another way of combining two vectors is with the use of the vector (or outer) product. A description of this product can be developed with reference to Fig. 5. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are two arbitrary vectors drawn from a common origin, they define a plane, providing of course that θ, the angle between them, lies in the range $0 < \theta < \pi$. If a vector $\boldsymbol{C}$ is constructed at the same origin and perpendicular to the plane, Eq. (12) leads to

$$\boldsymbol{C} \cdot \boldsymbol{A} = C_x A_x + C_y A_y + C_z A_z = 0, \tag{16}$$

and

$$\boldsymbol{C} \cdot \boldsymbol{B} = C_x B_x + C_y B_y + C_z B_z = 0. \tag{17}$$

Equations (16) and (17) form a pair of simultaneous, homogeneous equations. They cannot be solved uniquely for the components of $\boldsymbol{C}$. However, their solution can be found in terms of a parameter a. The result, which can be easily verified (problem 1), is

$$C_x = a(A_y B_z - A_z B_y), \tag{18}$$

$$C_y = a(A_z B_x - A_x B_z) \tag{19}$$

and

$$C_z = a(A_x B_y - A_y B_x). \tag{20}$$

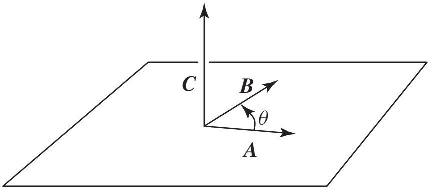


Fig. 5 The vector product $\boldsymbol{C} = \boldsymbol{A} \times \boldsymbol{B}$.

The parameter a is arbitrary and, for convenience, can be chosen equal to plus one. Then, from Eq. (1) and Eqs. (18) to (20),

$$\begin{aligned} C^2 = C_x^2 + C_y^2 + C_z^2 &= (A_x^2 + A_y^2 + A_z^2)(B_x^2 + B_y^2 + B_z^2) \\ &\quad - (A_x B_x + A_y B_y + A_z B_z)^2 \\ &= A^2 B^2 - A^2 B^2 \cos^2\theta \\ &= (AB \sin\theta)^2 \end{aligned} \tag{21}$$

Thus, the vector $\boldsymbol{C}$ represents the product of the vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ such that its length is given by $C = AB\sin\theta$. In the usual notation $\boldsymbol{C} = \boldsymbol{A} \times \boldsymbol{B}$.* This operation is referred to as the vector product of the two vectors and in the jargon used in this application it is called the "cross product". It must then be carefully distinguished from the dot product defined by Eq. (9).

With the use of Eqs. (18) to (20) the vector product can be written in the form

$$\boldsymbol{C} = \boldsymbol{A} \times \boldsymbol{B} = (A_y B_z - A_z B_y)\boldsymbol{i} + (A_z B_x - A_x B_z)\boldsymbol{j} + (A_x B_y - A_y B_x)\boldsymbol{k}, \tag{22}$$

which is represented by

$$\boldsymbol{A} \times \boldsymbol{B} = \boldsymbol{i} \begin{vmatrix} A_y & A_z \\ B_y & B_y \end{vmatrix} + \boldsymbol{j} \begin{vmatrix} A_x & A_z \\ B_x & B_z \end{vmatrix} + \boldsymbol{k} \begin{vmatrix} A_x & A_y \\ B_x & B_y \end{vmatrix}, \tag{23}$$

or more conveniently by the single determinant

$$\boldsymbol{A} \times \boldsymbol{B} = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \end{vmatrix}. \tag{24}$$

Following the general rules for the development of determinants (see Section 7.4), it is apparent that vector multiplication is not commutative, as $\boldsymbol{A} \times \boldsymbol{B} = -\boldsymbol{B} \times \boldsymbol{A}$. However, the normal distributive law still applies, as, for example,

$$\boldsymbol{A} \times (\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A} \times \boldsymbol{B} + \boldsymbol{A} \times \boldsymbol{C}. \tag{25}$$

From the definition of the vector product given above, it is clear that the magnitude of the vector $\boldsymbol{C}$ in Eq. (22) is equal to the area of the parallelogram defined by the vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ which describe its sides. However, there are two problems associated with this definition. First of all, the direction of the vector $\boldsymbol{C}$ is ambiguous in the absence of a convention. It is usually assumed, however, that the "right-hand rule" applies. Thus, if the first finger of the right

---

*The notation $\boldsymbol{C} = \boldsymbol{A} \wedge \boldsymbol{B}$ is used for the vector product in most texts in French.

hand is directed along $\boldsymbol{A}$ and the second along $\boldsymbol{B}$, the direction of the vector $\boldsymbol{C}$ is indicated by the thumb.

A second question arises for those who understand the importance of dimensional analysis, a subject that is treated briefly in Appendix II. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are both vector quantities with, say, dimensions of length, how can their cross product result in a vector $\boldsymbol{C}$, presumably with dimensions of length? The answer is hidden in the homogeneous equations developed above [Eqs. (18) to (20)]. The constant a was set equal to unity. However, in this case it has the dimension of reciprocal length. In other words, $C = aAB\,sin\,\theta$ is the length of the vector $\boldsymbol{C}$. In general, a vector such as **C** which represents the cross product of two "ordinary" vectors is an areal vector with different symmetry properties from those of $\boldsymbol{A}$ and $\boldsymbol{B}$.

4.5 TRIPLE PRODUCTS

Triple products involving vectors arise often in physical problems. One such product is $(\boldsymbol{A} \times \boldsymbol{B}) \times \boldsymbol{C}$, which is clearly represented by a vector. It is therefore called the vector triple product, whose development can be made as follows. If, in a Cartesian system, the vector $\boldsymbol{A}$ is chosen to be collinear with the x direction, $\boldsymbol{A} = A_x\boldsymbol{i}$. The vector $\boldsymbol{B}$ can, without loss of generality, be placed in the x,y plane. It is then given by $\boldsymbol{B} = B_x\boldsymbol{i} + B_y\boldsymbol{j}$. The vector $\boldsymbol{C}$ is then in a general direction, as given by $\boldsymbol{C} = C_x\boldsymbol{i} + C_y\boldsymbol{j} + C_z\boldsymbol{k}$, as shown in Fig. 6. Then, the cross products can be easily developed in the form $\boldsymbol{A} \times \boldsymbol{B} = A_xB_y\boldsymbol{k}$ and

$$(\boldsymbol{A} \times \boldsymbol{B}) \times \boldsymbol{C} = -A_xB_yC_y\boldsymbol{i} + A_xB_yC_x\boldsymbol{j}. \tag{26}$$

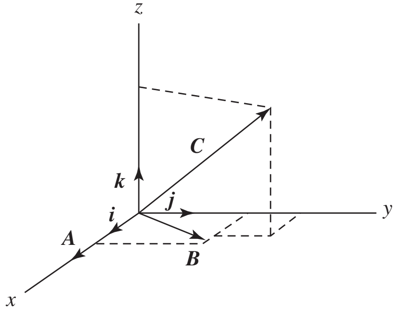


Fig. 6 Development of the triple product $(\boldsymbol{A} \times \boldsymbol{B}) \times \boldsymbol{C}$.

The evaluation of the scalar products $\boldsymbol{A}\cdot\boldsymbol{C}$ and $\boldsymbol{B}\cdot\boldsymbol{C}$ and substitution into Eq. (26) leads to the relation

$$(\boldsymbol{A}\times\boldsymbol{B})\times\boldsymbol{C}=(\boldsymbol{A}\cdot\boldsymbol{C})\boldsymbol{B}-(\boldsymbol{B}\cdot\boldsymbol{C})\boldsymbol{A}. \tag{27}$$

An analogous derivation can be carried out to obtain

$$\boldsymbol{C}\times(\boldsymbol{A}\times\boldsymbol{B})=(\boldsymbol{C}\cdot\boldsymbol{B})\boldsymbol{A}-(\boldsymbol{C}\cdot\boldsymbol{A})\boldsymbol{B} \tag{28}$$

(problem 4). The two expansions of the triple vector products given by Eqs. (27) and (28) are very useful in the manipulation of vector relations. Furthermore, vector multiplication is not associative. In general,

$$(\boldsymbol{A}\times\boldsymbol{B})\times\boldsymbol{C}\neq\boldsymbol{A}\times(\boldsymbol{B}\times\boldsymbol{C}), \tag{29}$$

as can be shown by developing Eqs. (27) and (28).

Consider now the vector product $\boldsymbol{A}\times\boldsymbol{B}$, where these vectors are shown in Fig. 7. It is perpendicular to the x,y plane and has a magnitude equal to A_xB_y, the area of the base of the parallelepiped. The height of the parallelepiped is given by $C_z=C\cos\theta$. Therefore, the volume of the parallelepiped is equal to $(\boldsymbol{A}\times\boldsymbol{B})\cdot\boldsymbol{C}=\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}$, which can also be written in the form of a determinant of the components, *viz.*

$$\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}=\begin{vmatrix}A_x & A_y & A_z\\ B_x & B_y & B_z\\ C_x & C_y & C_z\end{vmatrix}. \tag{30}$$

It should be noted that the positive sign of this result depends on the choice of a right-handed coordinate system in which the angle θ is acute. The relation developed here for the volume of a parallelepiped is often employed in crystallography to calculate the volume of a unit cell, as shown in the following section.

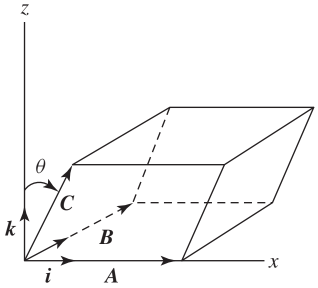


Fig. 7 Calculation of the volume of a parallelepiped.

An important symmetry property of the scalar triple product can be illustrated by the relations

$$(\boldsymbol{A} \times \boldsymbol{B}) \cdot \boldsymbol{C} = (\boldsymbol{B} \times \boldsymbol{C}) \cdot \boldsymbol{A} = (\boldsymbol{C} \times \boldsymbol{A}) \cdot \boldsymbol{B}, \tag{31}$$

that is, successive cyclic permutation. However, it changes sign upon interchange of any two vectors. These results follow directly from the properties of the determinant, Eq. (30). Furthermore, the value of the triple scalar product is not altered by the exchange of the symbols "dot" and "cross"; thus,

$$\boldsymbol{A} \times \boldsymbol{B} \cdot \boldsymbol{C} = \boldsymbol{B} \times \boldsymbol{C} \cdot \boldsymbol{A} = \boldsymbol{A} \cdot \boldsymbol{B} \times \boldsymbol{C}, \tag{32}$$

nor are the parentheses necessary in this case.

4.6 RECIPROCAL BASES

A set of three noncoplanar vectors forms a basis in a three-dimensional space. Any vector in this space can be represented by these three basis vectors. In certain applications, particularly in crystallography, it is convenient to define a second basis, in reciprocal space. Thus, if the vectors $\boldsymbol{t}_1, \boldsymbol{t}_2$ and $\boldsymbol{t}_3$ form a basis, in which $\boldsymbol{t}_1 \times \boldsymbol{t}_2 \cdot \boldsymbol{t}_3 \neq 0$, another basis can be defined by the vectors $\boldsymbol{b}_1, \boldsymbol{b}_2$ and $\boldsymbol{b}_3$. The two bases are said to be reciprocal if

$$\boldsymbol{t}_i \cdot \boldsymbol{b}_j = \delta_{\mathrm{i,j}}, \tag{33}$$

where $i, j = 1,2,3$ and $\delta_{i,j}$ is the Kronecker delta.* Thus, for example, $\boldsymbol{t}_1 \cdot \boldsymbol{b}_1 = 1, \boldsymbol{t}_1 \cdot \boldsymbol{b}_2 = 0$ and $\boldsymbol{t}_1 \cdot \boldsymbol{b}_3 = 0$. These relations show that $\boldsymbol{b}_1$ is perpendicular to both $\boldsymbol{t}_2$ and $\boldsymbol{t}_3$; it is therefore parallel to $\boldsymbol{t}_2 \times \boldsymbol{t}_3$. Then, $\boldsymbol{b}_1 = c\,\boldsymbol{t}_2 \times \boldsymbol{t}_3$, where c is a constant. Scalar multiplication by $\boldsymbol{t}_1$ then gives $\boldsymbol{t}_1 \cdot \boldsymbol{b}_1 = c\,\boldsymbol{t}_1 \cdot \boldsymbol{t}_2 \times \boldsymbol{t}_3 = 1$. These relations then lead to the expression

$$\boldsymbol{b}_1 = \frac{\boldsymbol{t}_2 \times \boldsymbol{t}_3}{\boldsymbol{t}_1 \cdot \boldsymbol{t}_2 \times \boldsymbol{t}_3}. \tag{34}$$

The corresponding relations for $\boldsymbol{b}_2$ and $\boldsymbol{b}_3$ follow by cyclic permutations of the subscripts (see Chapter 8).

An infinite three-dimensional crystal lattice is described by a primitive unit cell which generates the lattice by simple translations. The primitive cell can be represented by three basic lattice vectors such as $\boldsymbol{t}_1, \boldsymbol{t}_2$ and $\boldsymbol{t}_3$ defined above. They may or may not be mutually perpendicular, depending on the crystal

---

*Leopold Kronecker, German mathematician (1823–1891).

system. The volume of the primitive cell is equal to $\boldsymbol{t}_1 \times \boldsymbol{t}_2 \cdot \boldsymbol{t}_3$ and the position of each lattice point is specified by a vector

$$\boldsymbol{\tau}_n = n_1\boldsymbol{t}_1 + n_2\boldsymbol{t}_2 + n_3\boldsymbol{t}_3, \tag{35}$$

where n_1, n_2 and n_3 are integers.

The vectors which define the so-called reciprocal lattice are given by

$$\boldsymbol{k}_h = h_1\boldsymbol{b}_1 + h_2\boldsymbol{b}_2 + h_3\boldsymbol{b}_3, \tag{36}$$

where h_1, h_2 and h_3 are integers. The analog of the primitive cell in reciprocal space is known as the first Brillouin zone.* Its volume is given by

$$\boldsymbol{b}_1 \cdot \boldsymbol{b}_2 \times \boldsymbol{b}_3 = \frac{(\boldsymbol{t}_2 \times \boldsymbol{t}_3) \cdot [(\boldsymbol{t}_3 \times \boldsymbol{t}_1) \times (\boldsymbol{t}_1 \times \boldsymbol{t}_2)]}{(\boldsymbol{t}_1 \cdot \boldsymbol{t}_2 \times \boldsymbol{t}_3)^3} = \frac{1}{\boldsymbol{t}_1 \cdot \boldsymbol{t}_2 \times \boldsymbol{t}_3}\ . \tag{37}$$

The conclusion to be drawn from Eq. (37) is that the volume of the first Brillouin zone is equal to the reciprocal of the volume of the primitive cell. It should be noted that the scalar product

$$\boldsymbol{\tau}_n \cdot \boldsymbol{k}_h = n_1h_1 + n_2h_2 + n_3h_3 \tag{38}$$

is an integer.

4.7 DIFFERENTIATION OF VECTORS

If a vector $\boldsymbol{R}$ is a function of a single scalar quantity s, the curve traced as a function of s by its terminus, with respect to a fixed origin, can be represented as shown in Fig. 8. Within the interval $\Delta\boldsymbol{s}$ the vector $\Delta\boldsymbol{R} = \boldsymbol{R}_2 - \boldsymbol{R}_1$ is in the direction of the secant to the curve, which approaches the tangent in the limit as $\Delta s \to 0$. This argument corresponds to that presented in Section 2.3 and illustrated in Fig. 4 of that section. In terms of unit vectors in a Cartesian coordinate system

$$\boldsymbol{R} = \boldsymbol{i}\,R_x + \boldsymbol{j}\,R_y + \boldsymbol{k}\,R_z, \tag{39}$$

and

$$\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}t} = \boldsymbol{i}\frac{\mathrm{d}R_x}{\mathrm{d}t} + \boldsymbol{j}\frac{\mathrm{d}R_y}{\mathrm{d}t} + \boldsymbol{k}\frac{\mathrm{d}R_z}{\mathrm{d}t}. \tag{40}$$

Clearly, Eq. (40) includes variations in both the magnitude and direction of the vector $\boldsymbol{R}$. It is easily generalized to represent higher derivatives. For a function

---

*Léon Brillouin, French-American physicist (1889–1969).

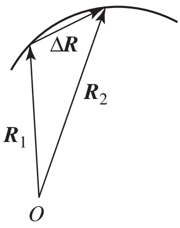


Fig. 8 Increment of a vector ***R***.

of two or more vectors, each of which depends on the single scalar parameter s, the usual rules of differentiation hold, as summarized in Section 2.3 for scalar quantities. However, the order of the vectors must not be changed in cases involving the vector product. Specifically, if $\boldsymbol{R}(s)$ and $\boldsymbol{S}(s)$ are differentiable vector functions,

$$\frac{\mathrm{d}(\boldsymbol{R}\times\boldsymbol{S})}{\mathrm{d}t}=\boldsymbol{R}\times\frac{\mathrm{d}\boldsymbol{S}}{\mathrm{d}t}+\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}t}\times\boldsymbol{S}, \tag{41}$$

where it is essential to preserve the order of the factors in each term on the right-hand side of Eq. (41).

4.8 SCALAR AND VECTOR FIELDS

The term scalar field is used to describe a region of space in which a scalar function is associated with each point. If there is a vector quantity specified at each point, the points and vectors constitute a vector field.

Suppose that $\phi(x,y,z)$ is a scalar point function, that is, a scalar function that is uniquely defined in a given region. Under a change of coordinate system to, say, x', y', z', it will take on another form, although its value at any point remains the same. Applying the chain rule (Section 2.12),

$$\frac{\partial\phi}{\partial x'}=\frac{\partial x}{\partial x'}\frac{\partial\phi}{\partial x}+\frac{\partial y}{\partial x'}\frac{\partial\phi}{\partial y}+\frac{\partial z}{\partial x'}\frac{\partial\phi}{\partial z}=a_{11}\frac{\partial\phi}{\partial x}+a_{12}\frac{\partial\phi}{\partial y}+a_{13}\frac{\partial\phi}{\partial z}, \tag{42}$$

$$\frac{\partial\phi}{\partial y'}=a_{21}\frac{\partial\phi}{\partial x}+a_{22}\frac{\partial\phi}{\partial y}+a_{23}\frac{\partial\phi}{\partial z} \tag{43}$$

and

$$\frac{\partial\phi}{\partial z'}=a_{31}\frac{\partial\phi}{\partial x}+a_{32}\frac{\partial\phi}{\partial y}+a_{33}\frac{\partial\phi}{\partial z}. \tag{44}$$

The quantities $\partial\phi/\partial x$, $\partial\phi/\partial y$ and $\partial\phi/\partial z$ are components of a vector,

$$\nabla\phi = \boldsymbol{i}\frac{\partial\phi}{\partial x} + \boldsymbol{j}\frac{\partial\phi}{\partial y} + \boldsymbol{k}\frac{\partial\phi}{\partial z}, \tag{45}$$

which has been transformed from one coordinate system to another. This operation can be written in more compact form with the use of matrix algebra, a subject that is developed in Chapter 7.

Equation (44) suggests that a vector operator ∇ or nabla (called "del") be defined in Cartesian coordinates by

$$\nabla = \boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}. \tag{46}$$

This operator is not a vector in the geometrical sense, as it has no scalar magnitude. However, it transforms as a vector and thus can be treated formally as such.

4.9 THE GRADIENT

The operator del is defined in Cartesian coordinates by Eq. (46). The result of its operation on a scalar is called the gradient. Thus, Eq. (45) is an expression for the gradient of ϕ, namely, $\nabla\phi = \boldsymbol{grad}\,\phi$, which is of course a vector quantity. The form of the differential operator del varies, however, depending on the choice of coordinates, as demonstrated in the following chapter.

To obtain a physical picture of the significance of the gradient, consider Fig. 9. The condition $d\phi = 0$ produces a family of surfaces such as that shown. The change in ϕ in passing from one surface to another will be the same regardless of the direction chosen. However, in the direction of $\boldsymbol{n}$, the normal

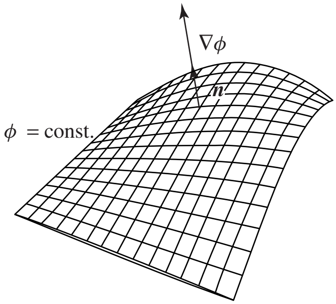


Fig. 9 The normal $\boldsymbol{n}$ to a surface and the gradient.

to the surface, the space rate of change of ϕ will be maximum. It is the change in ϕ in this direction that corresponds to the gradient.

4.10 THE DIVERGENCE

The scalar product of the vector operator ∇ and a vector $\boldsymbol{A}$ yields a scalar quantity, the divergence of $\boldsymbol{A}$. Thus,

$$\nabla \cdot \boldsymbol{A} = \boldsymbol{div}\,\boldsymbol{A} = \left[\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}\right] \cdot [\boldsymbol{i}A_x + \boldsymbol{j}A_y + \boldsymbol{k}A_z] \tag{47}$$

$$= \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}. \tag{48}$$

If $\boldsymbol{A}$ represents a vector field, the derivatives such as $\partial A_x/\partial x$ transform normally under a change of coordinates.

As a simple example of the divergence, consider the quantity $\nabla \cdot \boldsymbol{r}$, where $\boldsymbol{r} = \boldsymbol{x} + \boldsymbol{y} + \boldsymbol{z}$. Then,

$$\nabla \cdot \boldsymbol{r} = \left(\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}\right) \cdot (\boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}\boldsymbol{z})$$

$$= \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3. \tag{49}$$

4.11 THE CURL OR ROTATION

The vector product of ∇ and the vector $\boldsymbol{A}$ is known as the curl or rotation of $\boldsymbol{A}$. Thus in Cartesian coordinates,

$$\boldsymbol{curl}\,\boldsymbol{A} = \nabla \times \boldsymbol{A} = \boldsymbol{i}\left[\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right] + \boldsymbol{j}\left[\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right] + \boldsymbol{k}\left[\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right] \tag{50}$$

$$= \begin{vmatrix} \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ A_x & A_y & A_z \\ \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \end{vmatrix}. \tag{51}$$

In the development of the determinant in Eq. (51), care must be taken to preserve the correct order of the elements.

The following important relations involving the curl can be verified by expanding the vectors in terms of their components $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ in Cartesian coordinates:

$$\nabla \times (\boldsymbol{A} + \boldsymbol{B}) = \nabla \times \boldsymbol{A} + \nabla \times \boldsymbol{B}, \tag{52}$$

$$\nabla \times (\phi \boldsymbol{A}) = \nabla \phi \times \boldsymbol{A} + \phi \nabla \times \boldsymbol{A}, \tag{53}$$

$$\begin{aligned} \nabla(\boldsymbol{A} \cdot \boldsymbol{B}) = {} & (\boldsymbol{B} \cdot \nabla)\boldsymbol{A} + (\boldsymbol{A} \cdot \nabla)\boldsymbol{B} \\ & + \boldsymbol{B} \times (\nabla \times \mathbf{A}) + \boldsymbol{A} \times (\nabla \times \mathbf{B}), \end{aligned} \tag{54}$$

$$\nabla \cdot (\boldsymbol{A} \times \mathbf{B}) = \boldsymbol{B} \cdot \nabla \times \mathbf{A} - \boldsymbol{A} \cdot \nabla \times \boldsymbol{B} \tag{55}$$

and

$$\nabla \times (\boldsymbol{A} \times \mathbf{B}) = (\boldsymbol{B} \cdot \nabla)\boldsymbol{A} - \boldsymbol{B}(\nabla \cdot \mathbf{A}) - (\boldsymbol{A} \cdot \nabla)\boldsymbol{B} + \boldsymbol{A}(\nabla \cdot \mathbf{B}) \tag{56}$$

(problem 10).

4.12 THE LAPLACIAN*

In addition to the above vector relations involving del, there are six combinations in which del appears twice. The most important one, which involves a scalar, is

$$\nabla \cdot \nabla \phi = \nabla^2 \phi = \boldsymbol{div\,grad}\,\phi. \tag{57}$$

The operator ∇^2, which is known as the Laplacian, takes on a particularly simple form in Cartesian coordinates, namely,

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\ . \tag{58}$$

However, as shown in Section 5.15 it can become more complicated in other coordinate systems. When applied to a vector, it yields a vector, which is given in Cartesian coordinates by

$$\nabla^2 \boldsymbol{A} = \frac{\partial^2 \boldsymbol{A}}{\partial x^2} + \frac{\partial^2 \boldsymbol{A}}{\partial y^2} + \frac{\partial^2 \boldsymbol{A}}{\partial z^2}\ . \tag{59}$$

---

*Pierre Simon de Laplace, French astronomer and mathematician (1749–1827).

A third combination which involves del operating twice on a vector is

$$\nabla(\nabla\cdot\mathbf{A}) = \boldsymbol{grad\,div\,A} = \boldsymbol{i}\frac{\partial^2 A_x}{\partial x^2} + \boldsymbol{j}\frac{\partial^2 A_y}{\partial y^2} + \boldsymbol{k}\frac{\partial^2 A_z}{\partial z^2}$$
$$+\boldsymbol{i}\left[\frac{\partial^2 A_y}{\partial x\partial y} + \frac{\partial^2 A_z}{\partial x\partial z}\right] + \boldsymbol{j}\left[\frac{\partial^2 A_x}{\partial x\partial y} + \frac{\partial^2 A_z}{\partial y\partial z}\right] + \boldsymbol{k}\left[\frac{\partial^2 A_x}{\partial x\partial z} + \frac{\partial^2 A_y}{\partial y\partial z}\right]. \quad (60)$$

The cross product of two dels operating on a scalar function ϕ yields

$$\nabla\times\nabla\phi = \boldsymbol{curl\,grad}\,\phi$$
$$= \begin{vmatrix} \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \dfrac{\partial\phi}{\partial x} & \dfrac{\partial\phi}{\partial y} & \dfrac{\partial\phi}{\partial z} \\ \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \end{vmatrix} = 0. \quad (61)$$

If $\nabla\times\boldsymbol{A} = 0$ for any vector $\boldsymbol{A}$, then $\boldsymbol{A} = \nabla\phi$. In this case $\boldsymbol{A}$ is irrotational. Similarly,

$$\nabla\cdot\nabla\times\boldsymbol{A} = \boldsymbol{div\,curl\,A} = 0. \quad (62)$$

Finally, a useful expansion is given by the relation

$$\nabla\times(\nabla\times\boldsymbol{A}) = \boldsymbol{curl\,curl\,A} = \nabla(\nabla\cdot\boldsymbol{A}) - \nabla^2\boldsymbol{A}. \quad (63)$$

4.13 MAXWELL'S EQUATIONS

To illustrate the use of the vector operators described in the previous section, consider the equations of Maxwell. In a vacuum they provide the basic description of an electromagnetic field in terms of the vector quantities $\mathscr{E}$ the electric field and $\mathscr{H}$ the magnetic field. The definition of the field in a dielectric medium requires the introduction of two additional quantities, the electric displacement $\mathscr{D}$ and the magnetic induction $\mathscr{B}$. The macroscopic electromagnetic properties of the medium are then determined by Maxwell's equations, *viz.*

$$\nabla\cdot\mathscr{D} = \rho, \quad (64)$$

$$\nabla\cdot\mathscr{B} = 0, \quad (65)$$

$$\nabla\times\mathscr{H} = \boldsymbol{J} + \dot{\mathscr{D}} \quad (66)$$

and

$$\nabla\times\mathscr{E} = -\dot{\mathscr{B}}. \quad (67)$$

In these expressions ρ is the charge density in the medium and $\boldsymbol{J}$ is the current density.

In isotropic media $\mathscr{D}$ and $\mathscr{E}$ are related by $\mathscr{D} = \varepsilon\mathscr{E}$, where the scalar parameter ε is now referred to as the permittivity.* In the international (SI) system it is given by $\varepsilon = \varepsilon_r\varepsilon_0$, where ε_0 is the permittivity of vacuum (see Appendix II) and ε_r is a dimensionless permittivity that characterizes the medium. Furthermore, according to Ohm's law† the current is given by $\boldsymbol{J} = \sigma\mathscr{E}$, where σ is the electrical conductivity. The relation $\nabla \cdot \mathscr{B} = 0$ is a mathematical statement of the observation that isolated magnetic poles do not exist.

A very general relation, that is known as the equation of continuity, has applications in many branches of physics and chemistry. It can be derived by taking the divergence of Eq. (66). Then, from Eq. (62) the relation

$$\nabla \cdot \nabla \times \mathscr{H} = \nabla \cdot (\boldsymbol{J} + \dot{\mathscr{D}}) = \nabla \cdot \boldsymbol{J} + \dot{\rho} = 0, \tag{68}$$

and hence

$$\nabla \cdot \boldsymbol{J} = -\frac{\partial \rho}{\partial t}, \tag{69}$$

is easily obtained. This result can be interpreted in electromagnetic theory as follows. The divergence of the current density (flux) from a system must be compensated by the rate of decrease in charge density within the system. This statement is a special case of the general divergence theorem, which is derived in Appendix VI.

In atomic and molecular spectroscopy it is the electric field created by the light excitation that is the origin of the interaction with a sample. The effect of the magnetic field is several orders of magnitude weaker. In this application, then, unit relative permeability‡ will be assumed and $\mathscr{B}$ will be replaced by $\mu_0\mathscr{H}$ in Eqs. (65) and (67). Equations (64) to (67) become

$$\nabla \cdot \mathscr{D} = \rho, \tag{70}$$

$$\nabla \cdot \mathscr{H} = 0, \tag{71}$$

$$\nabla \times \mathscr{H} = \sigma\mathscr{E} + \boldsymbol{\varepsilon}\dot{\mathscr{E}} \tag{72}$$

and

$$\nabla \times \mathscr{E} = -\mu_0\dot{\mathscr{H}}, \tag{73}$$

---

*This quantity was previously called the dielectric constant. It is in general a function of frequency and therefore not a constant.

†Georg Simon Ohm, German Physicist (1789–1854).

‡In the SI system $\mathscr{B} = \mu\mathscr{H} = \mu_r\mu_0\mathscr{H}$, where μ is the permeability of the medium. Here again it is written as the product of the permeability of vacuum and a relative quantity μ_r, by analogy with the permittivity (see Appendix II).

respectively. The curl of Eq. (73) yields the relation

$$\nabla \times (\nabla \times \mathscr{E}) = -\mu_0 \nabla \times \dot{\mathscr{H}}, \tag{74}$$

$$= -\mu_0 \sigma \dot{\mathscr{E}} - \mu_0 \varepsilon \ddot{\mathscr{E}}, \tag{75}$$

where the time derivative of Eq. (72) has been substituted. With the use of the vector relation given by Eq. (63) the differential equation for the electric field can be written as

$$\nabla^2 \mathscr{E} - \nabla(\nabla \cdot \mathscr{E}) = \mu_0 \varepsilon \ddot{\mathscr{E}} + \mu_0 \sigma \dot{\mathscr{E}}. \tag{76}$$

It can be easily demonstrated that plane-wave solutions to Eq. (76) are of the form

$$\mathscr{E} = \mathscr{E}^0 e^{-2\pi i(\boldsymbol{k}\cdot\boldsymbol{r}+\nu t)} \tag{77}$$

for monochromatic waves of frequency ν propagating in the direction of $\boldsymbol{r}$ (see problem 14). Here, $\boldsymbol{k}$ is the propagation vector in reciprocal space.

From Eq. (77) the relations

$$\dot{\mathscr{E}} = -2\pi i \nu \mathscr{E}, \tag{78}$$

$$\nabla \cdot \mathscr{E} = -2\pi i \boldsymbol{k} \cdot \mathscr{E} \tag{79}$$

and

$$\nabla \times \mathscr{E} = -2\pi i \boldsymbol{k} \times \mathscr{E} \tag{80}$$

can be easily obtained. Their substitution in Eq. (76) yields

$$-(\boldsymbol{k} \cdot \mathscr{E})\boldsymbol{k} + (\boldsymbol{k} \cdot \boldsymbol{k})\mathscr{E} = -\nu^2 \mu_0 \left(\varepsilon + \frac{\sigma i}{2\pi \nu}\right) \mathscr{E} = \left(\frac{\nu}{c_0}\right)^2 \hat{n}^2 \mathscr{E}, \tag{81}$$

where by definition $\hat{n}^2 = \varepsilon_r + \sigma i/2\pi\nu\varepsilon_0$ is the square of the complex refractive index of the medium. By taking the scalar product of $\boldsymbol{k}$ with Eq. (81) it is found that

$$\hat{n}^2 \boldsymbol{k} \cdot \mathscr{E} = 0. \tag{82}$$

Thus, as $\hat{n}$ is not in general equal to zero, $\boldsymbol{k} \cdot \mathscr{E} = 0$, which describes a transverse wave, with the electric field perpendicular to the direction of propagation. The complex refractive index is then given by

$$\hat{n} = \sqrt{\varepsilon_r + \frac{\sigma i}{2\pi \nu \varepsilon_0}} = n + i\kappa, \tag{83}$$

the fundamental relation between the electrical and optical properties of a material. Note that in a nonconducting medium ($\sigma = 0$) the permittivity is equal to the square of the (real) refractive index.

In an anisotropic solid both ε and σ become tensor quantities, that is they are represented by 3×3 matrices (see Section 7.3). In general, then, a solid may exhibit anisotropy with respect to both the real and imaginary parts of the refractive index.

4.14 LINE INTEGRALS

Line integrals were introduced in Section 3.4.3. The principles presented there can be easily recast within the vector formalism of this chapter. Thus,

$$\int_c \boldsymbol{A} \cdot \mathrm{d}s \tag{84}$$

is one form of the line integral from a to b along curve 1, as shown in Fig. 10. Its evaluation, which results in a scalar quantity, can be carried out if $\boldsymbol{A} \cdot \mathrm{d}s$ is known as a function of the coordinates, say, x, y, z. A special case arises in which the function to be integrated is an exact differential (see Section 2.13). Thus, if

$$\boldsymbol{A} = \nabla\phi, \tag{85}$$

where ϕ is a scalar point function,

$$\begin{aligned}\int_a^b \boldsymbol{A} \cdot \mathrm{d}s = \int_a^b \nabla\phi \cdot \mathrm{d}s &= \int_a^b \left[\frac{\partial\phi}{\partial x}\mathrm{d}x + \frac{\partial\phi}{\partial y}\mathrm{d}y + \frac{\partial\phi}{\partial z}\mathrm{d}z\right] \\ &= \int_a^b \mathrm{d}\phi = \phi_b - \phi_a.\end{aligned} \tag{86}$$

If the integration is taken around a closed curve, as shown in Fig. 10,*

$$\int_a^a \nabla\phi \cdot \mathrm{d}s = \oint \nabla\phi \cdot \mathrm{d}s = 0. \tag{87}$$

Conversely, if $\oint \nabla\phi \cdot \mathrm{d}s = 0$, then Eq. (85) must hold and $\boldsymbol{A}$ is the gradient of some scalar point function ϕ. In conclusion, if $\boldsymbol{A} = \nabla\phi$, the line integral

---

*The symbol $\oint \mathrm{d}s$ represents a line integral around a closed path.

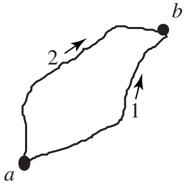


Fig. 10 Evaluation of line integrals.

$\int_a^b \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{s}$ depends only on the initial and final values of ϕ and is independent of the path.

The results obtained above are of fundamental importance in many physical problems. In mechanics, for example, a system is said to be conservative if the force on a given particle is given by

$$\boldsymbol{f} = -\nabla\phi, \tag{88}$$

where ϕ is a scalar potential function. Thus, from Eq. (61), $\nabla \times \boldsymbol{f} = 0$, and the force is irrotational. Furthermore, $\oint \nabla\phi \cdot \mathrm{d}\boldsymbol{s} = 0$, as shown.

In thermodynamics the state functions are independent of the path. That is, the reversible processes involved in passing from a given initial state to the final state are not involved in the resulting changes in such functions. The differentials of state functions are of course exact, as shown in Section 3.5.

4.15 CURVILINEAR COORDINATES

In previous sections of this chapter, vectors have been described by their components in a Cartesian system. However, for most physical problems it is not the most convenient one. It is generally important to choose a system of coordinates that is compatible with the natural symmetry of the problem at hand. This natural symmetry is determined by the boundary conditions imposed on the solutions.

If the Cartesian coordinates x, y and z are related to three new variables by

$$x = x(\xi_1, \xi_2, \xi_3), \tag{89}$$

$$y = y(\xi_1, \xi_2, \xi_3) \tag{90}$$

and

$$z = z(\xi_1, \xi_2, \xi_3). \tag{91}$$

The chain rule leads to expressions such as

$$\mathrm{d}x = \frac{\partial x}{\partial \xi_1}\,\mathrm{d}\xi_1 + \frac{\partial x}{\partial \xi_2}\,\mathrm{d}\xi_2 + \frac{\partial x}{\partial \xi_3}\,\mathrm{d}\xi_3, \tag{92}$$

with analogous relations for the other differentials. The most important case is that in which the new coordinates are orthogonal; that is, their surfaces $\xi_i = \text{constant}$, ($i = 1,2,3$) intersect at right angles. Then, the square of the distance between two adjacent points is given by

$$(\mathrm{d}s)^2 = (\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2 = h_1^2(\mathrm{d}\xi_1)^2 + h_2^2(\mathrm{d}\xi_2)^2 + h_3^2(\mathrm{d}\xi_3)^2, \tag{93}$$

where the h_i's are scale factors, with

$$h_i^2 = \left(\frac{\partial x}{\partial \xi_i}\right)^2 + \left(\frac{\partial y}{\partial \xi_i}\right)^2 + \left(\frac{\partial z}{\partial \xi_i}\right)^2. \qquad i = 1,2,3 \tag{94}$$

The distance between two points on a coordinate line is the line element

$$\mathrm{d}s_i = h_i\,\mathrm{d}\xi_i. \qquad i = 1,2,3 \tag{95}$$

Thus, the element of volume becomes equal to

$$\mathrm{d}s_1\,\mathrm{d}s_2\,\mathrm{d}s_3 = h_1 h_2 h_3\,\mathrm{d}\xi_1\,\mathrm{d}\xi_2\,\mathrm{d}\xi_3. \tag{96}$$

As explained in Section 5.9, each component of $\mathrm{d}\phi/\mathrm{d}s_i = (1/h_i)(\partial\phi/\partial\xi_i)$ is its directional derivative. In a curvilinear system its component perpendicular to the surface $\xi_i = \text{constant}$ (that is, in the direction of s_i) is

$$\frac{\mathrm{d}\phi}{\mathrm{d}s_i} = \frac{1}{h_i}\frac{\partial \phi}{\partial \xi_i}, \tag{97}$$

following Eq. (95). Then $\nabla\phi$ can be written in the form

$$\nabla\phi = \frac{\boldsymbol{e}_1}{h_1}\frac{\partial \phi}{\partial \xi_1} + \frac{\boldsymbol{e}_2}{h_2}\frac{\partial \phi}{\partial \xi_2} + \frac{\boldsymbol{e}_3}{h_3}\frac{\partial \phi}{\partial \xi_3}, \tag{98}$$

where the $\boldsymbol{e}_i$'s are unit vectors along the curvilinear coordinate axes.

It is now necessary to derive analogous relations for the divergence of a vector, *viz.* $\nabla \cdot \boldsymbol{A}$. The calculation can be carried out in at least two ways. The direct analytic approach is long, but does not involve any methods other than those of vector algebra. Otherwise, it is necessary to develop the divergence (Gauss's) theorem, after which the desired result is easily obtained (see Appendix VI). In either case it is given by

$$\nabla \cdot \boldsymbol{A} = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial \xi_1}(A_1 h_2 h_3) + \frac{\partial}{\partial \xi_2}(A_2 h_1 h_3) + \frac{\partial}{\partial \xi_3}(A_3 h_1 h_2)\right]. \tag{99}$$

If $\boldsymbol{A} = \nabla\phi$,

$$\nabla \cdot \nabla\phi = \nabla^2\phi = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial\xi_1}\left(\frac{h_2h_3}{h_1}\frac{\partial\phi}{\partial\xi_1}\right) + \frac{\partial}{\partial\xi_2}\left(\frac{h_1h_3}{h_2}\frac{\partial\phi}{\partial\xi_2}\right) + \frac{\partial}{\partial\xi_3}\left(\frac{h_1h_2}{h_3}\frac{\partial\phi}{\partial\xi_3}\right)\right] \tag{100}$$

as the components of $\nabla\phi$ are $A_i = 1/h_i(\partial\phi/\partial\xi_i)$ [see Eq. (98)]. Analogous expressions for $\nabla^2\boldsymbol{A}$ can be obtained with use of the expansion

$$\nabla^2\boldsymbol{A} = \nabla(\nabla \cdot \boldsymbol{A}) - \nabla \times \nabla \times \boldsymbol{A}. \tag{101}$$

The general expressions developed in this section can be applied to a given problem by calculating the h_i's from Eq. (94), providing of course that the coordinate transformations given by Eqs. (89) to (91) are known. Some well-known examples will be treated in the following chapter.

PROBLEMS

1. Show that Eqs. (16) and (17) are verified by the substitutions of Eqs. (18–20).

2. Given two vectors $\boldsymbol{A} = 4\boldsymbol{i} + \boldsymbol{j} + 3\boldsymbol{k}$ and $\boldsymbol{B} = \boldsymbol{i} - 3\boldsymbol{j} - \boldsymbol{k}$, calculate: $\boldsymbol{A} + \boldsymbol{B}$, $\boldsymbol{A} \cdot \boldsymbol{B}$ and $\boldsymbol{A} \times \boldsymbol{B}$. Ans. $5\boldsymbol{i} - 2\boldsymbol{j} + 2\boldsymbol{k}$, -2, $8\boldsymbol{i} + 7\boldsymbol{j} - 13\boldsymbol{k}$

3. If $\boldsymbol{A} = 2\boldsymbol{i} + 4\boldsymbol{j} + \boldsymbol{k}$ and $\boldsymbol{B} = -2\boldsymbol{i} + \boldsymbol{j} + 2\boldsymbol{k}$, find A, B, $\boldsymbol{A} \cdot \boldsymbol{B}$, and $cos\,\theta$. Ans. $\sqrt{21}$, 3, 2, 0.1455

4. Verify Eqs. (27) and (28).

5. Demonstrate the inequality of Eq. (29).

6. With the use of Eq. (24), calculate the volume of the parallelepiped defined by the vectors $\boldsymbol{A} = \boldsymbol{i} + 2\boldsymbol{j} + \mathbf{k}$, $\boldsymbol{B} = \boldsymbol{j} + \boldsymbol{k}$ and $\boldsymbol{C} = \boldsymbol{i} - \boldsymbol{j}$. Ans. 4

7. Show that $(\boldsymbol{A} \times \boldsymbol{B}) \cdot (\boldsymbol{C} \times \boldsymbol{D}) = (\boldsymbol{A} \cdot \boldsymbol{C})(\boldsymbol{B} \cdot \boldsymbol{D}) - (\boldsymbol{A} \cdot \boldsymbol{D})(\boldsymbol{B} \cdot \boldsymbol{C})$.

8. Calculate the angles between two diagonals of a cube. Ans. $cos^{-1}\frac{1}{3}$

9. Find the angle between the diagonal of a cube and a diagonal of a face. Ans. $cos^{-1}\sqrt{2/3}$

10. Demonstrate Eqs. (52–56) by expansion in Cartesian coordinates.

11. Show that if a vector $\boldsymbol{A}$ is irrotational, $\boldsymbol{A} = \nabla\phi$, where ϕ is a scalar.

12. Prove Eqs. (62) and (63).

13. Given a force $\boldsymbol{f} = \boldsymbol{i} - z\boldsymbol{j} - y\boldsymbol{k}$, show that it is conservative, *i.e.* that $\nabla \times \boldsymbol{f} = 0$. Find a scalar potential ϕ such that $\boldsymbol{f} = -\nabla\phi$. Ans. $\phi = -x + yz$.

14. Show that Eq. (77) represents a solution to Eq. (76).

15. From Eq. (3) derive the relations for the real and imaginary parts of the refractive index as functions of the permittivity and the electrical conductivity of a given medium. Note that both n and κ are defined as real quantities.

$$\text{Ans. } \kappa^2 = \frac{1}{2}\left[-\varepsilon_r + \sqrt{\varepsilon_r^2 + \frac{\sigma^2}{4\pi^2\nu^2\varepsilon_0^2}}\right]$$

$$\text{and } n^2 = \frac{1}{2}\left[\varepsilon_r + \sqrt{\varepsilon_r^2 + \frac{\sigma^2}{4\pi^2\nu^2\varepsilon_0^2}}\right],$$

where the positive square roots are to be taken.

15. Show that in the case of a relatively poor conductor, $\kappa \approx (\sigma/4\pi\nu\varepsilon_0)$ and $n \approx \sqrt{\varepsilon_r}$.

5 Ordinary Differential Equations

Differential equations are usually classified as "ordinary" or "partial". In the former case only one independent variable is involved and its differential is exact. Thus there is a relation between the dependent variable, say $y(x)$, its various derivatives, as well as functions of the independent variable x. Partial differential equations contain several independent variables, and hence partial derivatives.

The order of an ordinary differential equation is the order of its highest derivative. Thus, an ordinary differential equation of order n is an equation of the form

$$F(x, y, y', \dots , y^{(n)}) = 0. \tag{1}$$

If the dependent variable $y(x)$ and all of its derivatives occur in the first degree and do not appear as products, the equation is said to be linear. In effect, the solution of a differential equation of order n necessitates n integrations, each of which involves an arbitrary constant. However, in some cases one or more of these constants may be assigned specific values. The results, which are also solutions of the differential equation, are referred to as particular solutions. The general solution, however, includes all of the n constants of integration, whose evaluation requires additional information associated with the application.

5.1 FIRST-ORDER DIFFERENTIAL EQUATIONS

A first-order differential equation can always be solved, although its solution is not necessarily easy to obtain. If the variables are separable, the equation can be reduced to the form

$$f(x)\mathrm{d}x = g(y)\mathrm{d}y, \tag{2}$$

and the integration can usually be carried out by one of the methods illustrated in Section 3.3.

Furthermore, as shown in Section 3.5, a differential equation such as

$$N(x,y)\mathrm{d}x + M(x,y)\mathrm{d}y = 0 \tag{3}$$

can be integrated directly if the left-hand side is an exact differential. Although most differential equations of this type are not exact, in principle they can be made so by the introduction of a suitable integrating factor. If the equation is linear, which is often the case, it can be written as

$$\frac{\mathrm{d}y}{\mathrm{d}x} + yp(x) = q(x). \tag{4}$$

Now, if a function $\mu(x)$ is chosen so that

$$p(x) = \frac{\mu'(x)}{\mu(x)} = \frac{\mathrm{d}}{\mathrm{d}x}\, ln\, \mu(x), \tag{5}$$

this function is

$$\mu(x) = e^{\int p(x)\mathrm{d}x}, \tag{6}$$

and Eq. (4) becomes

$$\frac{\mathrm{d}y}{\mathrm{d}x} + y\frac{\mu'(x)}{\mu(x)} = q(x). \tag{7}$$

If both sides are multiplied by $\mu(x)$, Eq. (7) can be written as

$$\mu(x)y' + y\mu'(x) = \frac{\mathrm{d}}{\mathrm{d}x}[\mu(x)y] = \mu(x)q(x). \tag{8}$$

Thus,

$$\mu(x)y = \int \mu(x)q(x)\mathrm{d}x + C, \tag{9}$$

and the function $\mu(x) = e^{\int p(x)\mathrm{d}x}$ is the desired integrating factor.

As an example, consider the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} - yx = x. \tag{10}$$

By comparison with Eq. (4), $p = -x$, $q = x$ and the integrating factor is the Gaussian function $\mu = e^{-\frac{1}{2}x^2}$. With the introduction of this factor in Eq. (10),

$$e^{-\frac{1}{2}x^2}\frac{\mathrm{d}y}{\mathrm{d}x} - xe^{-\frac{1}{2}x^2}y = xe^{-\frac{1}{2}x^2} \tag{11}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}x}(e^{-\frac{1}{2}x^2}y) = xe^{-\frac{1}{2}x^2}. \tag{12}$$

The solution to Eq. (10) is then obtained from

$$e^{-\frac{1}{2}x^2}y = \int xe^{-\frac{1}{2}x^2}\mathrm{d}x = -e^{-\frac{1}{2}x^2} + C, \tag{13}$$

or, simply,

$$y = Ce^{\frac{1}{2}x^2} - 1, \tag{14}$$

as can be easily verified by substitution.

5.2 SECOND-ORDER DIFFERENTIAL EQUATIONS

Many second-order differential equations arise in physical problems. Fortunately, most of them can be cast into a relatively simple form, namely,

$$P(x)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + Q(x)\frac{\mathrm{d}y}{\mathrm{d}x} + R(x)y = 0, \tag{15}$$

where $P(x)$, $Q(x)$ and $R(x)$ are polynomials. As the right-hand side of Eq. (15) is equal to zero in this case, the equation is said to be homogeneous and the method of series solution can be applied. This method is illustrated as follows.

5.2.1 Series solution

The dependent variable $y(x)$ is written in a power series, *viz.*

$$y(x) = a_0 + a_1x + a_2x^2 + \cdots = \sum_n a_nx^n. \tag{16}$$

Successive differentiation yields

$$\frac{\mathrm{d}y}{\mathrm{d}x} = a_1 + 2a_2x + 3a_3x^2 + \cdots = \sum_n na_nx^{n-1} \tag{17}$$

and

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 2a_2 + 6a_3x + 12a_4x^2 \cdots = \sum_n n(n-1)a_nx^{n-2}. \tag{18}$$

The polynomial coefficients are of the form

$$P(x) = p_0 + p_1x + p_2x^2 \cdots\ , \tag{19}$$

$$Q(x) = q_0 + q_1x + q_2x^2 \cdots \tag{20}$$

and

$$R(x) = r_0 + r_1 x + r_2 x^2 \cdots . \tag{21}$$

The result of the substitution of Eqs. (16) to (21) into the differential equation [Eq. (15)] can be collected in powers of x. The constants, that is, the coefficients of x^0, lead to the relation

$$2a_2 p_0 + a_1 q_0 + a_0 r_0 = 0. \tag{22}$$

Thus,

$$a_2 = -\frac{a_1 q_0 + a_0 r_0}{2p_0}, \tag{23}$$

a function of the two coefficients a_0 and a_1. Equating the coefficients of x will yield an expression for a_3, namely

$$a_3 = \frac{1}{6p_0}\left\{\left[\frac{r_0(p_1+q_0)}{p_0} - r_1\right]a_0 + \left[\frac{q_0(p_1+q_0)}{p_0} - (q_1+r_0)\right]a_1\right\}, \tag{24}$$

where the expression for a_2 given by Eq. (23) has been employed. In principle, this procedure can be continued to obtain successive coefficients a_n as functions of only a_0 and a_1, two constants of integration.

An over-simplified example of this method is provided by the differential equation

$$\frac{d^2y}{dx^2} - y = 0. \tag{25}$$

Here, by comparison with Eq. (15) $P(x) = 1$, $Q(x) = 0$ and $R(x) = -1$; thus, all three coefficients in Eq. (15) are independent of x. The dependent variable $y(x)$ and its derivatives are developed as above [Eqs. (16) and (18)]. Substitution into Eq. (25) yields the relations $a_2 = a_0/2$, $a_3 = a_1/6$, *etc.*, which can be generalized in the form of a recursion formula for the coefficients,

$$a_{n+2} = \frac{a_n}{(n+1)(n+2)}. \tag{26}$$

A particular solution to Eq. (25) can be obtained by posing $a_0 = a_1 = 1$; then,

$$y_1 = 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 + \cdots = e^x, \tag{27}$$

where the identification of the series as the exponential has been made [see Eq. (1-10)]. It is easily verified by substitution that the exponential is a solution. However, it is also easy to show that the function $y_2 = e^{-x}$ is another solution to Eq. (25). As the ratio of these two solutions, $y_1/y_2 = e^{2x} \neq 0$, they are independent particular solutions. The general solution can then be written as

$$y = Ay_1 + By_2 = Ae^x + Be^{-x}, \tag{28}$$

where the constants of integration, A and B, are to be determined by the appropriate boundary conditions. From the definitions of the hyperbolic functions *sinh x* and *cosh x* [Eqs. (1-44) and (1-45)], it should be evident that the solution given by Eq. (28) can also be expressed in terms of these functions (see problem 3).

If in Eq. (15), $R = +1$, Eq. (25) becomes

$$\frac{d^2y}{dx^2} + y = 0, \tag{29}$$

and the particular solutions in this case are of the form $e^{\pm ix}$, as can be verified by substitution. It should be noted that the particular solutions are in this case periodic. The general solution

$$y(x) = Ae^{ix} + Be^{-ix} \tag{30}$$

can be expressed in terms of the functions *sin x* and *cos x* by application of Euler's relation [Eq. (1-32)]. Here again, the constants of integration are determined by the boundary conditions imposed on the general solution.

5.2.2. The classical harmonic oscillator

The example presented above will now be developed, as it is a problem which arises frequently in many applications. The vibrations of mechanical systems and oscillations in electrical circuits are illustrated by the following simple examples. The analogous subject of molecular vibrations is treated with the use of matrix algebra in Chapter 9.

Consider a physical pendulum, as represented in Fig. 1. A mass m is attached by a spring to a rigid support. The spring is characterized by a force

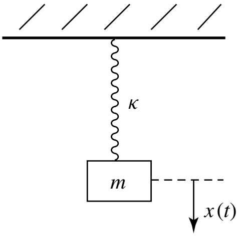


Fig. 1 Simple mechanical oscillator.

constant κ such that the force acting on the mass is described by Hooke's law,*

$$f = -\kappa x, \tag{31}$$

where $x(t)$ is the displacement of the mass from its equilibrium position and f is the force opposing this displacement (see Fig. 1).† Assuming that the force of gravity is independent of the small displacement $x(t)$, Newton's second law can be written in the form

$$f = m\ddot{x} = -\kappa x. \tag{32}$$

The equation of motion is then

$$\ddot{x} + \frac{\kappa}{m}x = 0. \tag{33}$$

In Eqs. (32) and (33) Newton's notation has been employed; the dot above a symbol indicates that its time derivative has been taken. Thus, $\mathrm{d}^2x/\mathrm{d}t^2 \equiv \ddot{x}$ is the second derivative of x with respect to time.

Aside from a constant and some changes in notation, Eq. (33) is of the same form as Eq. (29). Thus, particular solutions would be expected such as $e^{\pm i\omega_0 t}$, where $\omega_0 = 2\pi\nu^0$ is a constant and ν^0 is the natural frequency of oscillation. Substitution of this expression into Eq. (33) leads to the identification $\omega_0^2 = \kappa/m$. The general solution of Eq. (33) is then of the form

$$x(t) = Ae^{i\omega_0 t} + Be^{-i\omega_0 t}, \tag{34}$$

where A and B are two constants of integration. An alternative form of Eq. (34) is obtained from Euler's relation (Section 1.6), namely,

$$x = (A + B)\, cos\, \omega_0 t + i(A - B)\, sin\, \omega_0 t = C\, cos\, \omega_0 t + D\, sin\, \omega_0 t \tag{35}$$

and the constants C and D can also serve as the two integration constants.

Returning to the problem illustrated in Fig. 1, the question is: How is the pendulum put into motion at an initial time t_0?

(i) If at $t = t_0$ the mass is displaced by a distance x_0, and it is not given an initial velocity ($\dot{x}_0 = 0$), $C = x_0$ and $D = 0$. The solution is then given by

$$x = x_0\, cos\, \omega_0 t. \tag{36}$$

---

*Robert Hooke, English astronomer and mathematician (1635–1703).

†As shown in Section 5.14, in a conservative system the force can be represented by a potential function. The force is then given by $f = -\mathrm{d}V(x)/\mathrm{d}x$, where $V(x) = \frac{1}{2}\kappa x^2$ for this one-dimensional harmonic oscillator.

(ii) If at $t = t_0$ the mass is not displaced, but an initial velocity $\dot{x}_0 = v_0$ is imparted to it, as the derivative of Eq. (35) is

$$\dot{x} = -C\omega_0 \sin \omega_0 t + D\omega_0 \cos \omega_0 t, \tag{37}$$

$v_0 = D\omega_0$ and

$$x = \frac{v_0}{\omega_0} \sin \omega_0 t. \tag{38}$$

An alternative form of Eq. (35) can be obtained by substituting $C = \rho \cos \alpha$ and $D = \rho \sin \alpha$. Then,

$$x = \rho(\cos \alpha \cos \omega_0 t + \sin \alpha \sin \omega_0 t) = \rho \cos(\omega_0 t - \alpha). \tag{39}$$

The two constants of integration are now ρ and α, which are the amplitude and the phase angle, respectively. The initial conditions can be imposed as before.

5.2.3 The damped oscillator

Now suppose that the harmonic oscillator represented in Fig. 1 is immersed in a viscous medium. Equation (32) will then be modified to include a damping force which is usually assumed to be proportional to the velocity, $-h\dot{x}$. Thus, the equation of motion becomes

$$\ddot{x} + \frac{h}{m}\dot{x} + \frac{\kappa}{m}x = 0, \tag{40}$$

where the constant h depends on the viscosity of the medium.

The solution to Eq. (40) can be obtained with the substitution $x(t) = z(t)e^{\lambda t}$. The result is

$$e^{\lambda t}\left[\ddot{z} + \left(2\lambda + \frac{h}{m}\right)\dot{z} + \left(\lambda^2 + \frac{h\lambda}{m} + \frac{\kappa}{m}\right)z\right] = 0. \tag{41}$$

As the factor $e^{\lambda t} \neq 0$, the expression in brackets in Eq. (41) must be equal to zero. Furthermore, the parameter λ can be chosen so that the coefficient of $\dot{z}$ vanishes. Thus, $\lambda = -h/2m$ and Eq. (40) reduces to

$$\ddot{z} + \left(\frac{\kappa}{m} - \frac{h^2}{4m^2}\right)z = 0. \tag{42}$$

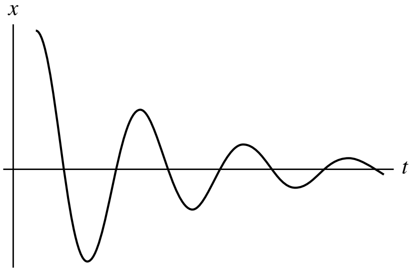


Fig. 2 Exponentially damped oscillation.

Here, two distinct situations arise depending on the relative magnitudes of the two terms in parentheses. If $\kappa/m > h^2/4m^2$, Eq. (42) is of the same form as Eq. (29), whose solutions can be written as $C\,cos\,\omega_1 t + D\,sin\,\omega_1 t$, with $\omega_1^2 = \kappa/m - h^2/4m^2$. Note that the presence of the damping term h/m modifies the natural (angular) frequency of the system. Then,

$$x = e^{-(h/2m)t}(C\,cos\,\omega_1 t + D\,sin\,\omega_1 t). \tag{43}$$

The two constants of integration, C and D, are determined as before by the initial conditions. This solution is oscillatory, although the amplitude of the oscillations decreases exponentially in time, as shown in Fig. 2.

On the other hand if $\kappa/m < h^2/4m^2$, the equation for $z(t)$ is of the form of Eq. (25) and the solutions are in terms of exponential functions of real arguments or hyperbolic functions. In this case $x(t)$ is not oscillatory and will simply decrease exponentially with time.

A third, very specific case occurs when $\kappa/m = h^2/4m^2$. The system is then said to be critically damped.

The mechanical problem treated above has its electrical analogy in the circuit shown in Fig. 3. It is composed of three elements, an inductance L, a capacitance C and a resistance R. If there are no other elements in the closed

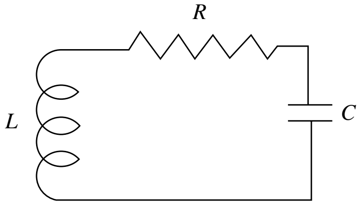


Fig. 3 Damped electrical oscillator.

circuit, according to Kirchhoff's second law,* the sum of the voltage drops across each of these three elements is equal to zero. The differential equation is then

$$L\frac{\mathrm{d}\iota}{\mathrm{d}t} + R\iota + \frac{q}{C} = 0, \tag{44}$$

where ι is the current and q is the charge on the capacitance. As the current is given by $\iota = \mathrm{d}q/\mathrm{d}t$, Eq. (44) becomes

$$\frac{\mathrm{d}^2q}{\mathrm{d}t^2} + \frac{R}{L}\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{1}{LC}q = 0, \tag{45}$$

which is of the same form as Eq. (40). Clearly, the resistance is responsible for the damping, while L and $1/C$ are analogous to the mass and force constant, respectively, which characterize the mechanical problem. This example will be treated in Chapter 11 with the use of the Laplace transform.

5.3 THE DIFFERENTIAL OPERATOR

The problems presented above can be solved with the use of an alternative method which employs operators of the type $\hat{\mathcal{D}} \equiv \mathrm{d}/\mathrm{d}x$. While the notion of operators will be developed in more detail in Chapter 7, it is sufficient here to point out that $\hat{\mathcal{D}}$ may be considered to be an abbreviation. This method can be applied in the case where $P(x)$, $Q(x)$ and $R(x)$ in Eq. (15) are constants, as in the examples considered above.

5.3.1 Harmonic oscillator

With the use of the differential operator the equation of motion for the harmonic oscillator [Eq. (29)], can be expressed as

$$(\hat{\mathcal{D}}^2 + 1)y = 0, \tag{46}$$

where the symbol $\hat{\mathcal{D}}^2$ is understood to mean two successive applications of the derivative. Formally, Eq. (46) can be factored, *viz.*

$$(\hat{\mathcal{D}} - r_1)(\hat{\mathcal{D}} - r_2)y = 0, \tag{47}$$

---

*Gustav Robert Kirchhoff, German physicist (1824–1887).

where r_1 and r_2 are the roots. Clearly, if y satisfies the equation

$$(\hat{\mathcal{D}} - r_1)y = 0, \tag{48}$$

its solution, $y = c_1 e^{r_1 x}$ is a particular solution of Eq. (46). An analogous argument for the second factor in Eq. (47) will lead to a second, independent particular solution of Eq. (46). The general solution is then of the form

$$y = c_1 e^{r_1 x} + c_2 e^{r_2 x}, \tag{49}$$

where both r_1 and r_2 are imaginary. With the changes in notation indicated above, this result is equivalent to Eq. (34) for the classical harmonic oscillator. This method can be easily extended to the example of the damped oscillator (see problem 7).

5.3.2 Inhomogeneous equations

If the right-hand side of Eq. (15) is not equal to zero, solutions are more difficult to obtain. Consider a second-order equation of the form

$$y'' + a_1 y' + a_2 y = f(x). \tag{50}$$

In terms of the differential operator it becomes

$$(\hat{\mathcal{D}}^2 + a_1 \hat{\mathcal{D}} + a_2)y = f(x), \tag{51}$$

or

$$(\hat{\mathcal{D}} - r_1)(\hat{\mathcal{D}} - r_2)y = f(x), \tag{52}$$

where r_1 and r_2 are the roots of the left-hand side of Eq. (51). It is convenient to make the substitution $u = (\hat{\mathcal{D}} - r_2)y$, which results in $(\hat{\mathcal{D}} - r_1)u = f(x)$, a linear first-order differential equation. It can be solved by application of the method outlined in Section 3.5. The integrating factor is then $exp(-\int r_1 \mathrm{d}x)$ and

$$u = e^{r_1 x} \int f(x) e^{-r_1 x} \mathrm{d}x + c_1 e^{r_1 x}. \tag{53}$$

The definition of u above, then leads to the relation

$$(\hat{\mathcal{D}} - r_2)y = e^{r_1 x}[g(x) + c_1], \tag{54}$$

where $g(x) = \int f(x) e^{-r_1 x} \mathrm{d}x$. Equation (54) can now be solved by the same procedure with the identification of $exp(-\int r_2 \mathrm{d}x)$ as the integrating factor. The result is

$$y = e^{r_2 x} \int g(x) e^{-(r_2 - r_1)x} \mathrm{d}x + \frac{c_1}{r_1 - r_2} e^{r_1 x} + c_2 e^{r_2 x}. \tag{55}$$

The coefficient in the second term on the right-hand side of Eq. (55) is a constant, so the sum of the second and third terms corresponds to the general solution of the homogeneous equation [Eq. (30)]. The first term is a particular integral which results from the nonzero term on the right-hand side of Eq. (50), *i.e.* the inhomogenuity. With the application of integration by parts, it can be written in the form

$$I = \frac{1}{r_1 - r_2}\left[e^{r_1 x}\int f(x)e^{-r_1 x}\mathrm{d}x - e^{r_2 x}\int f(x)e^{-r_2 x}\mathrm{d}x\right] \tag{56}$$

(see problem 8).

The reader is warned that the use of differential operators may lead to difficulties in certain cases. Specifically, if the coefficients appearing in Eq. (15) are functions of x, the method fails. Furthermore, it must be modified if two (or more) roots are equal.

5.3.3 Forced vibrations

An important example in mechanical and electrical systems is that of forced oscillations of a vibrational system. If an external force $f(t)$ is imposed on the mechanical oscillator considered above, Eq. (40) becomes

$$\ddot{x} + \frac{h}{m}\dot{x} + \frac{\kappa}{m}x = \frac{1}{m}f(t). \tag{57}$$

In practice, the right-hand side of Eq. (57) is often periodic in time, *e.g.* $f(t)/m = F_0\, sin\, \omega t$. The frequency ν of the applied force is equal to $\omega/2\pi$. Then, from Eq. (40) the inhomogeneous equation of interest is

$$\ddot{x} + \frac{h}{m}\dot{x} + \frac{\kappa}{m}x = F_0\, sin\, \omega t. \tag{58}$$

The general solution for the homogeneous part is given by Eq. (43) for the oscillatory (underdamped) case. The particular integral given by Eq. (56) can be developed as

$$I = \frac{1}{\left(\frac{\kappa}{m} - \omega^2\right)^2 + \frac{\omega^2 h^2}{m^2}}\left[\left(\frac{\kappa}{m} - \omega^2\right) sin\, \omega t - \frac{h\omega}{m}\, cos\, \omega t\right]. \tag{59}$$

The factor before the square brackets is of course the amplitude of the oscillation. It reaches a maximum value when the square of the angular frequency of the forcing function is given by

$$\omega^2 = \frac{\kappa}{m} - \frac{h^2}{2m^2}. \tag{60}$$

It should be noted that in the case of a damped oscillator, the condition given by Eq. (60) yields a resonant frequency that does not correspond to its natural frequency, as

$$\omega_1^2 = \omega_0^2 - \frac{h^2}{4m^2}. \tag{61}$$

The expression given by Eq. (59) is of particular importance in both mechanical and electrical systems. In the absence of damping, the amplitude of the forced oscillations approaches infinity at resonance. This result has been the origin of a number of well-known disasters, for example the collapse of the Tacoma Narrows bridge in the state of Washington in 1940. The turbulence created by strong winds in the narrow gorge produced periodic oscillations of the bridge which were, unfortunately, in resonance with the structure. A more classic example is that of the walls of Jericho that "came tumbling down", so it seems, because of resonance with the sound of the trumpets.

In electrical circuits the above analysis can be applied by adding an alternating voltage of angular frequency ω in series with the circuit shown in Fig. 3. However, the results in this case are normally less dramatic. In fact the condition of resonance, at which

$$\omega^2 = \frac{1}{LC} - \frac{R^2}{2L^2} \tag{62}$$

allows the resonant circuit of a radio receiver, for example, to be adjusted to correspond to the frequency of the detected signal. Usually, it is the capacitance, C, that is varied to achieve this condition.

5.4 APPLICATIONS IN QUANTUM MECHANICS

Most students are introduced to quantum mechanics with the study of the famous problem of the particle in a box. While this problem is introduced primarily for pedagogical reasons, it has nevertheless some important applications. In particular, it is the basis for the derivation of the translational partition function for a gas (Section 10.8.1) and is employed as a model for certain problems in solid-state physics.

5.4.1 The particle in a box

Consider a particle of mass m which is constrained to remain inside a one-dimensional "box" of width ℓ. The potential function which represents this system corresponds to

$$V(x) = \begin{cases} 0, & 0 < x < \ell \\ \infty, & x = 0, \ell \end{cases}. \tag{63}$$

In other words, there is no force acting on the particle except at the "walls" of the box. Schrödinger's second equation* for this problem (see Chapter 7) is then of the form

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V(x)\psi = \varepsilon\psi, \tag{64}$$

where, $\hbar \equiv h/2\pi$, h is Planck's constant† and ε is the energy of the single-particle system. The symbol ψ is by tradition used to represent the wavefunction, which describes the stationary (time-independent) states of the system.

In the interior of the box the particle is free; thus, $V(x) = 0$ and Eq. (64) becomes

$$\frac{d^2\psi}{dx^2} + \alpha^2\psi = 0, \tag{65}$$

where $\alpha^2 = 2m\varepsilon/\hbar^2$. Equation (65) is clearly of the same form (aside from notation) as Eq. (33). One form of its general solution is then

$$\psi(x) = \mathcal{A}\,sin(\alpha x + \eta), \tag{66}$$

by analogy with Eq. (39). The constant α can now be identified as $2\pi/\lambda$, where λ is the wavelength of a wave in the space of x. It is known in wave mechanics as the deBroglie‡ descriptive wave, with a wavelength given by

$$\lambda = \frac{2\pi}{\alpha} = \frac{\hbar}{\sqrt{2m\varepsilon}} = \frac{\hbar}{mv}. \tag{67}$$

In Eq. (67) the classical energy of a free particle, $\varepsilon = \frac{1}{2}mv^2$, has been substituted, with v its velocity and mv its momentum. Equation (67) is of course the well-known relation of deBroglie.

The solution of this problem, as given by Eq. (66), must now be analyzed with consideration of the boundary conditions at $x = 0$ and $x = \ell$. At these two points the potential function, $V(x)$, becomes infinite. Therefore, for the product $V(x)\ \psi(x)$ in Eq. (64) to remain finite at these two points, the wavefunction $\psi(x)$ must vanish. Clearly, if η, which is one of the arbitrary constants of integration, is chosen equal to zero in Eq. (66), the wavefunction will vanish at $x = 0$. However, at $x = \ell$ the situation is somewhat more complicated. A little reflection will show that if the argument of the sine

---

*Erwin Schrödinger, Austrian physicist (1887–1961).

†Max Planck, German physicist (1858–1947).

‡Louis deBroglie, French physicist (1892–1987).

function is equated to $n\pi x/\ell$, the wavefunction will vanish at $x = \ell$ for all values of the integer n. The acceptable solutions to this problem are then of the form

$$\psi_n(x) = \mathcal{A}\, sin\, \frac{n\pi x}{\ell}, \tag{68}$$

with $n = 1, 2, 3, \ldots$. The second constant of integration is the amplitude, $\mathcal{A}$, which is usually determined by normalizing $\psi(x)$.* Thus, the amplitude in Eq. (68) is determined by the condition that

$$\int_0^\ell |\psi_n(x)|^2 \mathrm{d}x = \mathcal{A}^2 \int_0^\ell sin^2\left(\frac{n\pi x}{\ell}\right) \mathrm{d}x = 1, \tag{69}$$

which yields $\mathcal{A} = \sqrt{2/\ell}$. The integral in Eq. (69) can be easily evaluated with the substitution $sin^2 y = \frac{1}{2}(1 - cos\, 2y)$. The wavefunctions for the first few values of n are represented in Fig. 4a.

With $\eta = 0$, the comparison of Eqs. (66) and (68) shows that $\alpha = n\pi/\ell$, and from Eq. (67) the energy is given by $\varepsilon = h^2n^2/8m\ell^2$, with $n = 1, 2, 3, \ldots$. Thus, the energy of the system is quantized due to the required boundary conditions on the solutions.

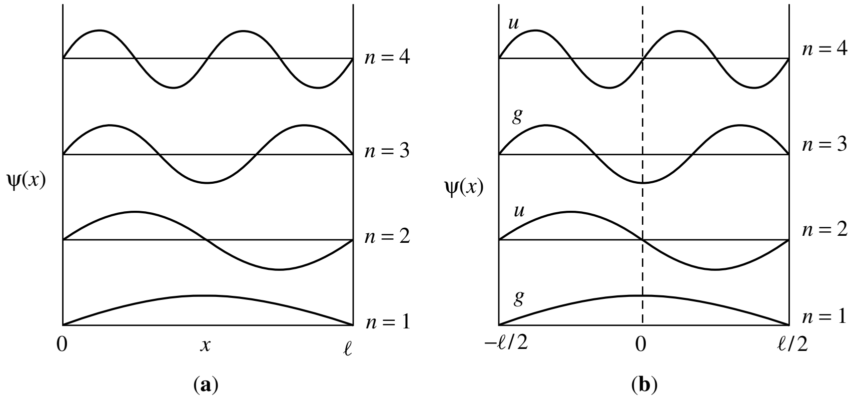


Fig. 4 Wavefunctions for the particle in a box: (**a**) without symmetry considerations; (**b**) the symmetric box.

---

*The normalization condition allows the quantity $\psi_v^*(\xi)\psi_v(\xi)\mathrm{d}\xi$ to be interpreted as the probability of finding the particle in the region of space $\mathrm{d}\xi$ (see Section 6.6.2).

5.4.2 Symmetric box

In the above treatment of the problem of the particle in a box, no consideration was given to its natural symmetry. As the potential function is symmetric with respect to the center of the box, it is intuitively obvious that this position should be chosen as the origin of the abscissa. In Fig. 4b, $x = 0$ at the center of the box and the walls are symmetrically placed at $x = \pm\ell/2$. Clearly, the analysis must in this case lead to the same result as above, because the particle does not "know" what coordinate system has been chosen! It is sufficient to replace x by $x + \ell/2$ in the solution given by Eq. (68). This operation is a simple translation of the abscissa, as explained in Section 1.2. The result is shown in Fig. 4b, where the wave function is now given by

$$\psi_n(x) = \mathcal{A}\, sin\left(\frac{n\pi x}{\ell} + \frac{n\pi}{2}\right). \tag{70}$$

It is easily verified that Eq. (70) satisfies the boundary conditions at the walls.

Although the net results obtained above for the particle in a box are physically the same, the mathematical consequences are quite different. From Fig. 4b it can be seen that the wavefunction is either even or odd, depending on the parity of n. Specifically, $\psi_n(x) = \pm\psi_n(-x)$, where the plus sign is appropriate when n is odd and the minus sign when n is even. As Eq. (70) contains the sine of the sum of two terms, it can be rewritten in the form

$$\psi_n(x) = \mathcal{A}\left(sin\,\frac{n\pi x}{\ell}\,cos\,\frac{n\pi}{2} + cos\,\frac{n\pi x}{\ell}\,sin\,\frac{n\pi}{2}\right); \tag{71}$$

then,

$$\psi_n^{(g)}(x) = \pm\mathcal{A}\, cos\,\frac{n\pi x}{\ell} \qquad \text{if } n \text{ is odd} \tag{72}$$

and

$$\psi_n^{(u)}(x) = \pm\mathcal{A}\, sin\,\frac{n\pi x}{\ell} \qquad \text{if } n \text{ is even.} \tag{73}$$

In spectroscopic applications the letters g and u (German: *Gerade, Ungerade*) are used to specify the symmetry of the functions under the inversion operation, $x \to -x$. Note that the normalization constant is given by $\mathcal{A} = \sqrt{2/\ell}$, as before.

The symmetry properties of the wavefunctions, as given by Eqs. (72) and (73) are extremely useful in the evaluation of certain integrals arising in quantum mechanics. First of all, it is evident that

$$\int_{-\ell/2}^{+\ell/2} \psi_n^{(g)}(x)\psi_n^{(u)}, \mathrm{d}x = 0 \tag{74}$$

for all values of n and n'. Other integrals of the type

$$\int_{-\ell/2}^{+\ell/2} \psi_n(x) f(x) \psi_{n'}(x) \mathrm{d}x \tag{75}$$

depend on the overall symmetry of the triple product in the integrand of Eq. (75). If the integrand is of symmetry "u", the integral is equal to zero. Clearly, the relations $g \times g = g, u \times u = g$ and $g \times u = u$ are applicable. These principles, which are the bases for the determination of spectroscopic selection rules, are developed in Sections 8.10 and 12.3.

5.4.3 Rectangular barrier: The tunnel effect

A relatively simple problem which has a direct application in the theory of chemical reaction rates is that of the rectangular barrier. A particle of mass m and energy $\varepsilon < V'$ approaches the barrier of height V' from the left (Fig. 5). Before the encounter with the barrier the amplitude of the deBroglie wave is equal to A, and after reflection by the barrier it is B. The wavefunction in region ①, where $x < 0$, is then $\psi_{①} = Ae^{i\alpha x} + Be^{-i\alpha x}$. The solution is periodic in this region, as $V = 0$ and $\alpha^2 = 2m\varepsilon/\hbar^2 > 0$. In region ②, with $\varepsilon < V'$, the solution is exponential, *viz.* $\psi_{②} = Ce^{\beta x} + De^{-\beta x}$, where $\beta^2 = 2m(V' - \varepsilon)/\hbar^2 > 0$. To the right of the barrier the solution is once again periodic, because $V = 0$, and the wavefunction is of the form $\psi_{③} = Fe^{i\alpha x}$, if it is assumed that the particle is not reflected at $x = \infty$.

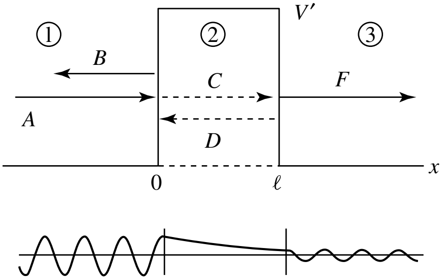


Fig. 5 Particle with a rectangular barrier.

At each boundary, $x = 0$ and $x = \ell$, both the function and its derivative must be continuous. These conditions impose the following relations:

$$\psi_{①}(0) = \psi_{②}(0), \left[\frac{\mathrm{d}\psi_{①}}{\mathrm{d}x}\right]_{x=0} = \left[\frac{\mathrm{d}\psi_{②}}{\mathrm{d}x}\right]_{x=0} \tag{76}$$

and

$$\psi_{②}(\ell) = \psi_{③}(\ell), \left[\frac{\mathrm{d}\psi_{②}}{\mathrm{d}x}\right]_{x=\ell} = \left[\frac{\mathrm{d}\psi_{③}}{\mathrm{d}x}\right]_{x=\ell}. \tag{77}$$

The application of Eqs. (76) and (77) to the solutions indicated above results in a system of four simultaneous equations:

(i) At $x = 0$,

$$\begin{aligned} A + B &= C + D \\ i\alpha A - i\alpha B &= \beta C - \beta D, \end{aligned} \tag{78}$$

and

(ii) At $x = \ell$,

$$\begin{aligned} Ce^{\beta\ell} + De^{-\beta\ell} &= Fe^{\beta\ell} \\ \beta Ce^{\beta\ell} - \beta De^{-\beta\ell} &= \alpha Fe^{\beta\ell}. \end{aligned} \tag{79}$$

As these functions cannot be normalized, it is sufficient here to pose $|A|^2 = 1$ and calculate the relative probability densities in each succeeding step. Then, $R = |B|^2$ represents the reflection coefficient and $T = |F|^2$ the transmission coefficient. Assuming that the particle cannot remain trapped within the barrier, the relation

$$|B|^2 + |F|^2 = 1 \tag{80}$$

represents the conservation of probability density in the system [see Eq. (69)].

After a bit of algebra it is found that the transmission coefficient is given by

$$T = \frac{1}{cosh^2\beta\ell + \dfrac{1}{4}\left(\dfrac{\alpha}{\beta} - \dfrac{\beta}{\alpha}\right)^2 sinh^2\beta\ell}. \tag{81}$$

Equation (81) can be verified by calculation of $R = |B|^2$ from the simultaneous equations for the coefficients and substitution in Eq. (80). The result represented by Eq. (81) shows that the transmission coefficient decreases as the height V' or the thickness ℓ of the barrier increases.

The possibility that a particle with energy less than the barrier height can penetrate is a quantum-mechanical phenomenon known as the tunnel effect. A number of examples are known in physics and chemistry. The problem illustrated here with a rectangular barrier was used by Eyring* to estimate the rates of chemical reactions. It forms the basis of what is known as the absolute reaction-rate theory. Another, more recent example is the inversion of the ammonia molecule, which was exploited in the ammonia maser – the forerunner of the laser (see Section 9.4.1).

5.4.4 The harmonic oscillator in quantum mechanics

One of the most important second-order, homogeneous differential equations is that of Hermite.† It arises in the quantum mechanical treatment of the harmonic oscillator. Schrödinger's equation for the harmonic oscillator leads to the differential equation

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}\xi^2} + (\sigma - \xi^2)\psi = 0, \tag{82}$$

where $\psi(\xi)$ is the wavefunction and σ is a constant. As a first step in the solution of this problem, it is useful to look for what is called the asymptotic solution, that is, the solution to Eq. (82) in the limit as $\xi^2 \to \infty$. Since in this case $\sigma \ll \xi^2$, Eq. (82) reduces to

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}\xi^2} - \xi^2\psi = 0, \tag{83}$$

with approximate solutions of the form $\psi(\xi) \approx Ce^{\pm\frac{1}{2}\xi^2}$. This function can be tested by consideration of its second derivative

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}\xi^2} = Ce^{\pm\frac{1}{2}\xi^2}(\xi^2 \pm 1) \approx C\xi^2 e^{\pm\frac{1}{2}\xi^2}. \tag{84}$$

This asymptotic solution suggests that the substitution $\psi(\xi) = \mathcal{H}_v(\xi)e^{\pm\frac{1}{2}\xi^2}$ in Eq. (82) should be tried. If the resulting differential equation for $\mathcal{H}(\xi)$ can be solved, the expression for $\psi(\xi)$ might be valid for all values of the independent variable ξ.

---

*Henry Eyring, American physical chemist (1901–1981).

†Charles Hermite, French mathematician (1822–1901).

The substitution proposed above leads to the well-known equation of Hermite,

$$\frac{d^2\mathcal{H}}{d\xi^2} - 2\xi\frac{d\mathcal{H}}{d\xi} + (\sigma - 1)\mathcal{H} = 0. \tag{85}$$

This equation can be solved by the method described in Section 5.2.1. The dependent variable is developed in a power series,

$$\mathcal{H}(\xi) = \sum_n a_n\xi^n, \tag{86}$$

by analogy with Eq. (16). Its first and second derivatives are found by term-by-term differentiation [see Eqs. (17) and (18)]. The substitution of these results in Eq. (85) leads directly to the expression

$$\sum_n n(n-1)a_n\xi^{n-2} - 2\sum_n na_n\xi^n + (\sigma - 1)\sum_n a_n\xi^n = 0. \tag{87}$$

It must be emphasized that the indices n appearing in each summation in Eq. (87) are independent. Thus, to collect the coefficients of, say, ξ^n, the index in the first term can be advanced, independently of the indices in the second and third terms. If in the first term n is replaced by $n+2$, it becomes $\sum_n (n+1)(n+2)a_{n+2}\xi^n$. Then, Eq. (87) can be written as a function of a single index, namely,

$$\sum_n \left[(n+1)(n+2)a_{n+2} - 2na_n + (\sigma - 1)a_n\right]\xi^n = 0. \tag{88}$$

Clearly, for this sum to vanish for all values of ξ, the coefficient in brackets must vanish for all values of n. Thus,

$$(n+1)(n+2)a_{n+2} - 2na_n + (\sigma - 1)a_n = 0 \tag{89}$$

and

$$a_{n+2} = \frac{2n - \sigma + 1}{(n+1)(n+2)}a_n. \tag{90}$$

This result is the recursion formula which allows the coefficient a_{n+2} to be calculated from the coefficient a_n. Starting with either a_0 or a_1 an infinite series can be constructed which is even or odd, respectively. These two coefficients are of course the two arbitrary constants in the general solution of a second-order differential equation. If one of them, say, a_0 is set equal to zero, the remaining series will contain the constant a_1 and be composed only of odd powers of ξ. On the other hand, if $a_1 = 0$, the even series will result. It can be shown, however, that neither of these infinite series can be accepted as solutions to the harmonic oscillator problem in quantum mechanics, as they are not convergent for large values of ξ.

The problem of convergence of the infinite series developed above can be circumvented by stopping the chosen series after a given finite number of terms. To break off the series at the term where $n = v$, it is sufficient to replace n by v in Eq. (90) and pose $2v - \sigma + 1 = 0$. The coefficient a_{v+2} then vanishes, yielding a polynomial of degree v. These functions are known as the Hermite polynomials. The factor $e^{\pm\frac{1}{2}\xi^2}$ introduced above will assure the required convergence if the negative sign is chosen in the exponent. The solution to Eq. (82) is then of the form $\psi(\xi) = \mathcal{H}_v(\xi)e^{-\frac{1}{2}\xi^2}$, where $\mathcal{H}_v(\xi)$ is the Hermite polynomial of degree v.

In the quantum mechanical application, the constant σ, is proportional to ε, the energy of the oscillator; namely,

$$\sigma = \frac{2\varepsilon}{h\nu^0}, \tag{91}$$

where h is Planck's constant and ν^0 is the frequency of the classical oscillator (see Section 5.2.2). The condition applied above, *viz.* $2v - \sigma + 1 = 0$ then leads to the well-known result

$$\varepsilon = h\nu^0\left(v + \tfrac{1}{2}\right), \tag{92}$$

where $v = 0, 1, 2, \ldots$, identified here as the degree of the Hermite polynomial. It is known to spectroscopists as the vibrational quantum number. It should be emphasized that this quantization of the energy is not determined by the differential equation in question, but by the condition imposed to assure the acceptability of its solution.

5.5 SPECIAL FUNCTIONS

The Hermite polynomials introduced above represent an example of special functions which arise as solutions to various second-order differential equations. After a summary of some of the properties of these polynomials, a brief description of a few others will be presented. The choice is based on their importance in certain problems in physics and chemistry.

5.5.1 Hermite polynomials

While the Hermite polynomials can be developed with the use of the recursion formula [Eq. (90)], it is more convenient to employ one of their fundamental definitions, *e.g.*

$$\mathcal{H}_v(\xi) \equiv (-1)^v e^{\xi^2} \frac{\mathrm{d}^v e^{-\xi^2}}{\mathrm{d}\xi^v}. \tag{93}$$

An alternative definition involves the use of a generating function. This method is especially convenient for the evaluation of certain integrals of the Hermite polynomials and can be applied to other polynomials as well. For the Hermite polynomials the generating function can be written as

$$S(\xi, s) \equiv e^{\xi^2-(s-\xi)^2} \equiv \sum_{v=0}^{\infty} \frac{\mathcal{H}_v(\xi)}{v!} s^v. \tag{94}$$

The variable s is a dummy variable in the sense that it does not enter the final result. Thus, if the exponential function in Eq. (94) is expanded in a power series in s, the coefficients of successive powers of s are just the Hermite polynomials $\mathcal{H}_v(\xi)$ divided by $v!$. It is not too difficult to show that Eqs. (93) and (94) are equivalent definitions of the Hermite polynomials.

Certain relations between the Hermite polynomials and their derivatives can be obtained from Eq. (94). First, the partial derivative of Eq. (94) with respect to s is

$$\frac{\partial S}{\partial s} = -2(s-\xi)S = \sum_{v=1}^{\infty} \frac{\mathcal{H}_v(\xi)}{v!} v s^{v-1} \tag{95}$$

and

$$-2(s-\xi) \sum_{v=0}^{\infty} \frac{\mathcal{H}_v(\xi)}{v!} s^v = \sum_{v=1}^{\infty} \frac{\mathcal{H}_v(\xi)}{(v-1)!} s^{v-1}. \tag{96}$$

By collecting the coefficients of a given power of s,

$$\sum_{v=0}^{\infty} \left[\frac{\mathcal{H}_{v+1}(\xi)}{v!} + \frac{2\mathcal{H}_{v-1}(\xi)}{(v-1)!} - \frac{2\xi\mathcal{H}_v(\xi)}{v!} \right] s^v = 0 \tag{97}$$

As this relation is correct for all values of s, the coefficients in brackets must vanish. The result yields the important recursion formula for the Hermite polynomials,

$$\mathcal{H}_{v+1}(\xi) - 2\xi\mathcal{H}_v(\xi) + 2v\mathcal{H}_{v-1}(\xi) = 0, \quad v = 1, 2, 3, \ldots \tag{98}$$

which is usually written in the form

$$\xi\mathcal{H}_v(\xi) = \tfrac{1}{2}\mathcal{H}_{v+1}(\xi) + v\mathcal{H}_{v-1}(\xi). \tag{99}$$

This relation can also be derived from the definition given by Eq. (93), which represents the series

$$\mathcal{H}_v(\xi) = (2\xi)^v - \frac{v(v-1)}{1!}(2\xi)^{v-2} + \frac{v(v-1)(v-2)(v-3)}{2!}(2\xi)^{v-4} - \cdots. \tag{100}$$

It breaks off at $(2\xi)^0$ or $(2\xi)^1$, depending on the parity of v. Differentiation of Eq. (100) leads to the expressions

$$\frac{\mathrm{d}\mathcal{H}_v(\xi)}{\mathrm{d}\xi} = 2v\mathcal{H}_{v-1}(\xi) \tag{101}$$

and

$$\frac{\mathrm{d}^2\mathcal{H}_v(\xi)}{\mathrm{d}\xi^2} = 2v\frac{\mathrm{d}\mathcal{H}_{v-1}(\xi)}{\mathrm{d}\xi} = 4v(v-1)\mathcal{H}_{v-2}(\xi). \tag{102}$$

Clearly, expressions for higher derivatives can be derived by the same method. Substitution of Eqs. (101) and (102) into Hermite's equation [Eq. (85)], with $\sigma - 1$ replaced by $2v$, leads to Eq. (99) (see problems 15 and 16).

In quantum mechanics it is customary to multiply the wavefunctions introduced in Eq. (82) by a normalizing factor, $\mathcal{N}_v$. Then,

$$\psi_v(\xi) = \mathcal{N}_v e^{-\frac{1}{2}\xi^2}\mathcal{H}_v(\xi) \tag{103}$$

and these functions form an orthonormal set for all values of ξ such that

$$\int_{-\infty}^{\infty} \psi_{v'}^*(\xi)\psi_v(\xi)\mathrm{d}\xi = \delta_{v',v} = \begin{cases} 1 & \text{if} \quad v' = v \\ 0 & \text{if} \quad v' \neq v \end{cases}, \tag{104}$$

where the symbol $\delta_{v',v}$ is known as the delta of Kronecker. If the $v' \neq v$, the integral in Eq. (104) is equal to zero and the functions are orthogonal. On the other hand, if $v' = v$, it is equal to one and the functions are normal – hence the term "orthonormal". This geometrical interpretation is derived from vector analysis (see Section 4.3).

Now take $v' < v$ and consider the integral

$$I = \int_{-\infty}^{\infty} \mathcal{H}_{v'}(\xi)\mathcal{H}_v(\xi)e^{-\xi^2}\mathrm{d}\xi = (-1)^v \int_{-\infty}^{\infty} \mathcal{H}_{v'}(\xi)\frac{\mathrm{d}^v(e^{-\xi^2})}{\mathrm{d}\xi^v}\mathrm{d}\xi. \tag{105}$$

Integration by parts (see Section 3.3.2) yields

$$I = (-1)^v \left[\mathcal{H}_{v'}(\xi)\frac{\mathrm{d}^{v-1}(e^{-\xi^2})}{\mathrm{d}\xi^{v-1}}\right]_{-\infty}^{\infty} - (-1)^v \int_{-\infty}^{\infty} \frac{\mathrm{d}\mathcal{H}_{v'}(\xi)}{\mathrm{d}\xi}\frac{\mathrm{d}^{v-1}(e^{-\xi^2})}{\mathrm{d}\xi^{v-1}}\mathrm{d}\xi. \tag{106}$$

The first term on the right-hand side of Eq. (106) vanishes, as the Gaussian function and its derivatives are equal to zero at $\xi = \pm\infty$. From Eq. (101) $\mathrm{d}\mathcal{H}_{v'}(\xi)/\mathrm{d}\xi = 2v'\mathcal{H}_{v'-1}(\xi)$ and Eq. (106) becomes

$$I = 2v'(-1)^{v+1} \int_{-\infty}^{\infty} \mathcal{H}_{v'-1}(\xi) \frac{\mathrm{d}^{v-1}(e^{-\xi^2})}{\mathrm{d}\xi^{v-1}} \mathrm{d}\xi. \tag{107}$$

If this process is continued, the result is

$$I = 2^{v'}(-1)^{v+v'} v'! \int_{-\infty}^{\infty} \mathcal{H}_0(\xi) \frac{\mathrm{d}^{v-v'}(e^{-\xi^2})}{\mathrm{d}\xi^{v-v'}} \mathrm{d}\xi, \tag{108}$$

$$= 2^{v'}(-1)^{v+v'} v'! \left[\frac{\mathrm{d}^{v-v'-1}(e^{-\xi^2})}{\mathrm{d}\xi^{v-v'-1}} \right]_{-\infty}^{\infty} = 0. \tag{109}$$

If $v = v'$, Eq. (105) becomes

$$\int_{-\infty}^{\infty} [\mathcal{H}_v(\xi)]^2 e^{-\xi^2} \mathrm{d}\xi = 2^v(-1)^{2v} v! \int_{-\infty}^{\infty} e^{-\xi^2} \mathrm{d}\xi = 2^v v! \sqrt{\pi} \tag{110}$$

and Eq. (104) is verified if the normalizing factor is taken to be

$$\mathcal{N}_v = \frac{1}{\sqrt{2^v v! \sqrt{\pi}}}. \tag{111}$$

Some of the Hermite polynomials and the corresponding harmonic-oscillator wave functions are presented in Table 1. The importance of the parity of these functions under the inversion operation, $\xi \to -\xi$ cannot be overemphasized.

5.5.2 Associated Legendre* polynomials

As shown in Chapter 6, these functions arise in all central-force problems, that is, systems composed of two interacting spherical objects in free space. The fundamental differential equation involved is

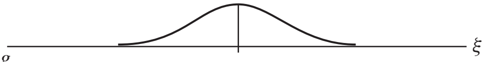
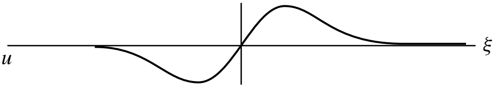
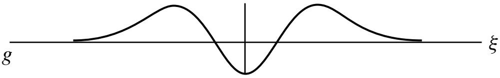
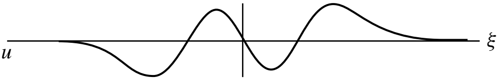
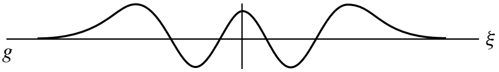
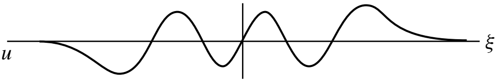
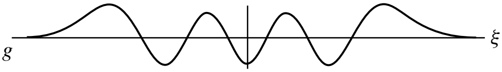
$$(1 - z^2) \frac{\mathrm{d}^2 P(z)}{\mathrm{d}z^2} - 2z \frac{\mathrm{d}P(z)}{\mathrm{d}z} + \left(\beta - \frac{m^2}{1 - z^2} \right) P(z) = 0, \tag{112}$$

where β is a constant and $m = 0, \pm 1, \pm 2, \ldots$ (see Section 6.4.2). If m is equal to zero, this equation can be solved by the development of $P(z)$ in

---

*Adrien-Marie Le Gendre, French mathematician (1752–1833).

Table 1 The Hermite polynomials and the harmonic-oscillator wavefunctions.

$\mathcal{H}_v(\xi)$	Symmetry $\psi_v = \mathcal{N}_v e^{-\frac{1}{2}\xi^2}\mathcal{H}_v(\xi)$
$\mathcal{H}_0(\xi) = 1$	g 
$\mathcal{H}_1(\xi) = 2\xi$	u 
$\mathcal{H}_2(\xi) = 4\xi^2 - 2$	g 
$\mathcal{H}_3(\xi) = 8\xi^3 - 12\xi$	u 
$\mathcal{H}_4(\xi) = 16\xi^4 - 48\xi^2 + 12$	g 
$\mathcal{H}_5(\xi) = 32\xi^5 - 160\xi^3 + 120\xi$	u 
$\mathcal{H}_6(\xi) = 64\xi^6 - 480\xi^4 + 720\xi^2$	g 

a power series, as before. However, if $m \neq 0$, the problem becomes more difficult due to the presence of the term with $(1 - z^2)$ in the denominator. At the points where $z = \pm 1$ this term becomes infinite. At these points, which are called singular points, the method of integration in series usually breaks down. However, if these points correspond to nonessential singularities (or regular points), it is often possible to avoid this problem with the use of appropriate substitutions. Here, with

$$P(z) = (1 - z^2)^s G(z) \tag{113}$$

the so-called index $s \geq 0$ is determined by inserting Eq. (113) in Eq. (112). The resulting terms in $(1 - z^2)^{s-1} = (1 - z^2)^s / 1 - z^2$ are

$$\frac{4z^2 s(s-1) + 4z^2 s - m^2}{1 - z^2} = \frac{4z^2 s^2 - m^2}{1 - z^2} = -m^2. \tag{114}$$

With a little reflection it can be seen that the second equality results if s is chosen so that $4s^2 = m^2$ or $s = \pm m/2$. Thus the troublesome factor $(1 - z^2)^{-1}$

has been eliminated. Furthermore, the condition that $s \geq 0$ then imposes the result $s = |m|/2$ and Eq. (113) becomes

$$P(z) = (1 - z^2)^{|m|/2} G(z). \tag{115}$$

The differential equation for $G(z)$ is

$$(1 - z^2)\frac{\mathrm{d}^2 G(z)}{\mathrm{d}z^2} - 2z(1 + |m|)\frac{\mathrm{d}G(z)}{\mathrm{d}z} + [\beta - |m|(|m| + 1)]G(z) = 0, \tag{116}$$

which can be solved directly by the series method.

The substitution $G(z) = \sum_n b_n z^n$ results in the relation

$$\sum_n n(n-1)b_n z^{n-2} - \sum_n n(n-1)b_n z^n - 2\sum_n (1 + |m|)n b_n z^n + \sum_n [\beta - |m|(|m| + 1)]b_n z^n = 0. \tag{117}$$

Here again the indices n are independent in each summation, so that n can be replaced by $n + 2$ in the first term. Then, by posing the coefficient of z^n equal to zero, the recursion formula becomes

$$b_{n+2} = \frac{(n + |m|)(n + |m| + 1) - \beta}{(n + 1)(n + 2)} b_n. \tag{118}$$

Once again there is a problem of convergence, this time at the points $z = \pm 1$. It is therefore necessary to break off the series at the term $n = n'$, where

$$\beta = (n' + |m|)(n' + |m| + 1) = \ell(\ell + 1). \tag{119}$$

The new integer $\ell = n' + |m| = |m|, |m| + 1, |m| + 2, \ldots$ is therefore related to m by the condition $|m| \leq \ell$ or

$$m = 0, \pm 1, \pm 2, \ldots \pm \ell. \tag{120}$$

It will be identified in Chapter 6 as the azimuthal quantum number, which is characteristic of the two-body problem.

The associated Legendre polynomials can be defined by the generating function

$$T_{|m|}(z, t) \equiv \frac{(2|m|)!(1 - z^2)^{|m|/2} t^{|m|}}{2^{|m|}(|m|)!(1 - 2zt + t^2)^{|m|+\frac{1}{2}}} \equiv \sum_{\ell=|m|}^{\infty} P_\ell^{|m|}(z) t^\ell. \tag{121}$$

It is analogous to the generating function for the Hermite polynomials [Eq. (94)], although somewhat more complicated. It can be used to obtain the useful recursion relations

$$zP_\ell^{|m|}(z) = \frac{(\ell + |m|)}{(2\ell + 1)} P_{\ell-1}^{|m|}(z) + \frac{(\ell - |m| + 1)}{(2\ell + 1)} P_{\ell+1}^{|m|}(z), \tag{122}$$

$$(1 - z^2)^{\frac{1}{2}} P_\ell^{|m|-1}(z) = \frac{1}{(2\ell + 1)} P_{\ell+1}^{|m|}(z) - \frac{1}{(2\ell + 1)} P_{\ell-1}^{|m|}(z) \tag{123}$$

and

$$\begin{aligned}(1 - z^2)^{\frac{1}{2}} P_\ell^{|m|+1}(z) = {} & \frac{(\ell + |m|)(\ell + |m| + 1)}{(2\ell + 1)} P_{\ell-1}^{|m|}(z) \\ & - \frac{(\ell - |m|)(\ell - |m| + 1)}{(2\ell + 1)} P_{\ell+1}^{|m|}(z)\end{aligned} \tag{124}$$

(see problem 20).

An alternative definition, but equally useful, of the associated Legendre polynomials is of the form

$$P_\ell^m(z) = \frac{(1 - z^2)^{m/2}}{2^n \ell!} \frac{\mathrm{d}^{\ell+m}}{\mathrm{d}z^{\ell+m}} (z^2 - 1)^\ell. \tag{125}$$

It is analogous to the definition of the Hermite polynomials, as given by Eq. (93).

When the associated Legendre polynomials are normalized they are written in the form

$$\Theta_{\ell,m}(\theta) = \sqrt{\frac{(2\ell + 1)}{2} \frac{(\ell - |m|)!}{(\ell + |m|)!}} P_\ell^{|m|}(cos\,\theta), \tag{126}$$

where the independent variable z has been replaced by $cos\,\theta$ and the normalizing constant has been evaluated by much the same procedure as that employed for the Hermite polynomials. The functions $\Theta_{\ell,m}(\theta)$ form an orthonormal set in the sense that

$$\int_0^\pi \Theta_{\ell',m}(\theta)\Theta_{\ell,m}(\theta)\, sin\,\theta\, \mathrm{d}\theta = \delta_{\ell',\ell}. \tag{127}$$

The explicit form of the normalized associated Legendre polynomials is given by

$$\Theta_{\ell,m}(\theta) = \frac{(-1)^\ell}{2^\ell \ell!} \sqrt{\frac{2\ell + 1}{2} \frac{(\ell - |m|)!}{(\ell + |m|)!}} sin^{|m|}\theta \frac{\mathrm{d}^{\ell+|\mathrm{m}|}(sin^{2\ell}\theta)}{(\mathrm{d}\, cos\,\theta)^{\ell+|m|}}. \tag{128}$$

They often appear as products of the function $1/\sqrt{2\pi}e^{im\varphi}$. The angles θ and φ are just the two angles defined in spherical coordinates, as shown in Fig. 6-5. The function *sin* θ appearing in the integral arises from the appropriate volume element. The functions

$$Y_\ell^m(\theta, \varphi) = \Theta_{\ell,m}(\theta)\frac{1}{\sqrt{2\pi}}e^{im\varphi} \tag{129}$$

are known as spherical harmonics (see Appendix III).

5.5.3 The associated Laguerre polynomials*

Consider the differential equation

$$\frac{\mathrm{d}^2R(\rho)}{\mathrm{d}\rho^2} + \frac{2}{\rho}\frac{\mathrm{d}R(\rho)}{\mathrm{d}\rho} + \left[\frac{\gamma}{\rho} - \frac{\ell(\ell+1)}{\rho^2} - \frac{1}{4}\right]R(\rho) = 0, \tag{130}$$

where γ is a constant and $\ell = 0, 1, 2 \ldots$. As in the problem of the harmonic oscillator (Section 4.4.4), it is of interest to discuss first the asymptotic solution as $\rho \to \infty$. In this limit the terms in $1/\rho$ approach zero and Eq. (130) becomes

$$\frac{\mathrm{d}^2R(\rho)}{\mathrm{d}\rho^2} - \frac{1}{4}R(\rho) = 0. \tag{131}$$

Particular solutions to Eq. (131) are $R(\rho) = e^{\pm\rho/2}$, where only the negative exponent yields an acceptable function at infinity. This result suggests the substitution $R(\rho) = e^{-\rho/2}S(\rho)$, which results in the differential equation

$$\frac{\mathrm{d}^2S(\rho)}{\mathrm{d}\rho^2} + \left(\frac{2}{\rho} - 1\right)\frac{\mathrm{d}S(\rho)}{\mathrm{d}\rho} + \left[\frac{\gamma - 1}{\rho} - \frac{\ell(\ell+1)}{\rho^2}\right]S(\rho) = 0. \tag{132}$$

This equation cannot be solved by expansion in series, as the coefficients of $S(\rho)$ and its first derivative result in a singularity at $\rho = 0$. Because this point is regular, the substitution $S(\rho) = \rho^s\mathcal{L}(\rho)$ is suggested. If the coefficient of ρ^{s-2} is set equal to zero, the resulting indicial equation is

$$2s + s(s-1) - \ell(\ell+1) = 0. \tag{133}$$

Its solutions are

$$s = \ell, -\ell - 1. \tag{134}$$

The second solution in Eq. (134) is not compatible with the condition $s \geq 0$. Therefore, the substitution $S(\rho) = \rho^\ell\mathcal{L}(\rho)$ is introduced into Eq. (130), leading

---

*Edmond Laguerre, French mathematician (1834–1886).

to the differential equation

$$\rho \frac{\mathrm{d}^2\mathcal{L}(\rho)}{\mathrm{d}\rho^2} + [2(\ell+1) - \rho]\frac{\mathrm{d}\mathcal{L}(\rho)}{\mathrm{d}\rho} + (\gamma - \ell - 1)\mathcal{L}(\rho) = 0. \tag{135}$$

This equation is of the form of Eq. (15) and hence can be solved by the power-series expansion $\mathcal{L}(\rho) = \sum_k a_k \rho^k$. The resulting recursion formula is

$$a_{k+1} = \frac{k + \ell + 1 - \gamma}{k(k+1) + 2(\ell+1)(k+1)} a_k. \tag{136}$$

Unlike the previous two examples, this is a one-term recursion formula. Hence, the series that is constructed from the value of a_0 is a particular solution of Eq. (135). Once again, however, because of the problem of convergence, the series must be terminated after a finite number of terms. The condition for it to break off after the term in $\rho^{k'}$ is given by

$$k' + \ell + 1 - \gamma = 0. \tag{137}$$

As the integers k' and ℓ both begin at zero, $\gamma = 1, 2, 3 \ldots$ can of course be identified as the principal quantum number n for the hydrogen atom (see Section 6.6.1). Thus, the quantization of the energy is due to the termination of the series, a condition imposed to obtain an acceptable solution. The associated Laguerre polynomials provide quantitative descriptions of the radial part of the wave functions for the hydrogen atom, as described in Appendix IV.

5.5.4 The gamma function

The gamma function is a generalization of the factorial introduced in Section 1.4. There, the notation $n! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot \ldots n$ was employed, with n a positive integer (or zero). The gamma function in this case is chosen so that $\Gamma(n) = (n-1)!$. However, a general definition due to Euler states that

$$\Gamma(z) = \lim_{n\to\infty} \frac{1 \cdot 2 \cdot 3 \cdots (n-1)}{z(z+1)\cdots(z+n-1)} n^z. \tag{138}$$

Several properties of the gamma function follow from this definition, *e.g.*

$$\Gamma(z+1) = z\Gamma(z), \tag{139}$$

$$\Gamma(1) = \lim_{n\to\infty} \frac{n!}{n!} = 1 \tag{140}$$

and, if n is a positive integer,

$$\Gamma(n) = (n-1)! \tag{141}$$

as stated above. It is also apparent that from the definition given by Eq. (138) that $\Gamma(z)$ becomes infinite at $z = 0, -1, -2, \ldots$, but is continuous (analytic) everywhere else.

An alternative expression for the gamma function is

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} \mathrm{d}t, \tag{142}$$

which is valid when the real part of z is positive. The evaluation of some of the gamma functions give $\Gamma(0) = \infty$, $\Gamma(1) = 1$, $\Gamma(2) = 1$, $\Gamma(3) = 2!$, $\Gamma(4) = 3!$, *etc.*. Furthermore, if $\Gamma(z)$ is known for $0 < z < 1$, $\Gamma(z)$ can be calculated for all real, positive values of z with the use of Eq. (139). Finally, for half-integer values of the argument, starting with $z = \frac{1}{2}$, Eq. (142) becomes

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty e^{-t} t^{-1/2} \mathrm{d}t = 2 \int_0^\infty e^{-x^2} \mathrm{d}x = \sqrt{\pi} \tag{143}$$

and similarly, $\Gamma(\frac{3}{2}) = \frac{1}{2}\sqrt{\pi}$, $\Gamma(\frac{5}{2}) = \frac{3}{4}\sqrt{\pi}$, $\Gamma(\frac{7}{2}) = \frac{15}{8}\sqrt{\pi}$, *etc.*

5.5.5 Bessel functions*

Bessel's equation can be written in the form

$$x^2 y'' + x y' + (x^2 - k^2) y = 0, \tag{144}$$

where k is a constant. The substitution $y(x) = x^\ell$ leads to the indicial equation $\ell^2 = k^2$. The roots are then $\pm k$. A particular solution is of the form

$$y = J_k(x) = \sum_{\lambda=0}^{\infty} \frac{(-1)^\lambda}{\Gamma(\lambda+1)\Gamma(\lambda+k+1)} \left(\frac{x}{2}\right)^{k+2\lambda}, \tag{145}$$

where $J_k(x)$ is the Bessel function of order k. It can be shown that if the difference between the values of the two roots $\pm k$ obtained above is not an integer, the general solution is given by

$$y(x) = A J_k(x) + B J_{-k}(x). \tag{146}$$

Even in the case where $k = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \cdots$ a general solution of the type given in Eq. (146) can be found. In fact, this case is of particular importance in many physical problems, as these Bessel functions are closely related to the ordinary trigonometric functions.

---

*Friedrich Wilhelm Bessel, German astronomer (1784–1846).

To illustrate this relationship, substitute $y = ux^{-\frac{1}{2}}$ in Eq. (144). The result is another form of Bessel's equation, namely,

$$u'' + \left(1 - \frac{4p^2 - 1}{4x^2}\right) u = 0. \tag{147}$$

In the special case in which $p = \pm\frac{1}{2}$ Eq. (147) reduces to

$$\frac{d^2u}{dx^2} + u = 0, \tag{148}$$

whose solution is sinusoidal [see Eq. (30)]. More generally, if p is finite, Eq. (147) becomes Eq. (148) in the limit as $x \to \infty$. Specifically, the Bessel functions of half-integer order are then given by

$$\lim_{x\to\infty} J_{n+\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}}\, sin(x - \tfrac{1}{2}n\pi), \tag{149}$$

where n is an integer. The corresponding functions of negative order are often referred to as Neumann functions.* Certain linear combinations of the Bessel and Neumann functions are known as Hankel functions.† The reader is referred to advanced texts for the various recurrence relations among these functions, as well as their integral representations.

5.5.6 Mathieu functions‡

These functions arise in a certain number of problems in electromagnetic theory and acoustics – in particular, those involving the vibrations of elliptical drum heads and the waves on approximately elliptical lakes. For the physical chemist, their interest is primarily in the treatment of the problem of internal rotation in a molecule. For example, the methyl group, CH_3, can assume three equivalent minimal positions around the single bond with which it is attached to the rest of a molecule (see Section 9.4.2). In general, if α represents the angle of internal rotation, the potential function for the rotation of a given functional group can be written in a first approximation in the form

$$V(\alpha) = \frac{V_N}{2}(1 - cos\, N\alpha). \tag{150}$$

---

*Johann (John) von Neumann, American mathematician (1903–1957).

†Hermann Hankel, German mathematician (1839–1873).

‡Emile Léonard Mathieu, French mathematician (1835–1890).

Here N represents the order of the rotation axis, *i.e.* $N = 3$ for the hindered rotation of a methyl group about its C_3 symmetry axis (see Chapter 9).

The Schrödinger equation for the hindered rotator can be written in the form

$$\frac{\hbar^2}{2I}\frac{d^2\psi(\alpha)}{d\alpha^2} + \left[\varepsilon - \frac{V_N}{2}(1 - cos\ N\alpha)\right]\psi(\alpha) = 0, \tag{151}$$

where I is the moment of inertia of the rotator* and ε is the energy. Comparison of Eq. (151) with the general form of Mathieu's equation,

$$\frac{d^2y}{dx^2} + (a - 16b\,cos\,2x)y = 0, \tag{152}$$

yields the relations: $y = \psi(\alpha), x = N\alpha/2$,

$$a = \frac{8I(\varepsilon - \frac{1}{2}V_N)}{\hbar^2N^2} \text{ and } b = -\frac{IV_N}{4\hbar^2N^2}.$$

Although Eq. (152) can in principle be solved by the development of $y(x)$ in a power series, the periodicity of the argument of cosine, namely, $2x = N\alpha$ complicates the problem. The most important application of Mathieu's equation to internal rotation in molecules is in the analysis of the microwave spectra of gases and vapors. The needed solutions to equations such as Eq. (152) are usually obtained numerically.

5.5.7 The hypergeometric functions

A differential equation due to Gauss is of the form

$$x(x-1)\frac{d^2y}{dx^2} + [(1+\alpha+\beta)x - \gamma]\frac{dy}{dx} + \alpha\beta y = 0, \tag{153}$$

where α, β and γ are constants. Substitution of a power series, $y = \sum_{n=0}^{\infty} a_n x^n$ leads to the one-term recursion formula

$$a_{n+1} = \frac{(\alpha+n)(\beta+n)}{(n+\gamma)(n+1)}a_n. \tag{154}$$

The resulting series is a particular solution to Eq. (153) known as the hypergeometric series. It converges for $|x| < 1$. It is usually denoted as $F(\alpha,\beta;\gamma,x)$.

---

*Strictly speaking, I is the reduced moment of inertia for the relative rotational motion of the system. For the case of a relatively light rotor such as CH_3 it is the moment of inertia of the hindered rotor that appears in Eq. (143).

Although the hypergeometric functions are useful in spectroscopy, as they describe the rotation of a symmetric top molecule (Section 9.2.4), their importance is primarily due to their generality. If, for example, $\alpha = 1$ and $\beta = \gamma$, Eq. (154) becomes $a_{n+1} = a_n$ for all values of n. The result is the ordinary geometric series

$$y = 1 + x + x^2 + x^3 \dots . \tag{155}$$

If the substitution $x = \frac{1}{2}(1 - z)$ is made in Eq. (153), the result is the differential equation of Legendre, with $\alpha = \ell + 1$, $\beta = -1$ and $\gamma = 1$ [see Eq. (112) with $m = 0$].

The Chebyshev polynomials,* which occur in quantum chemistry and in certain numerical applications, can be obtained from the hypergeometric functions by placing $\alpha = -\beta$, an integer, and $\gamma = \frac{1}{2}$. Finally, the hypergeometric functions reduce to the Jacobi polynomials† of degree n if $n = -\alpha$ is a positive integer.

PROBLEMS

1. Verify that $y = Ce^{\frac{1}{2}x^2} - 1$ is a solution to Eq. (10).
2. Derive Eq. (24).
3. Express Eq. (28) in terms of hyperbolic functions.
   Ans. $y = (A + B)\,cosh\,x + (A - B)\,sinh\,x$
4. Verify that Eq. (30) is one form of the general solution to Eq. (29).
5. Verify Eqs. (35) and (39).
6. Show that the two particular solutions proposed for Eq. (46) are independent.
7. Solve Eq. (45) with the use of the operator $\hat{\mathcal{D}} = \mathrm{d}/\mathrm{d}t$ and find the condition for critical damping. Ans. $R = 2\sqrt{L/C}$
8. Verify Eqs. (55) and (56).
9. Derive Eq. (59), verify Eq. (60) and show that Eq. (61) expresses the resonance condition.
10. With the use of Eqs. (66) and (68) show that the energy of the particle in the box is given by $\varepsilon = hn^2/8m\ell^2$, with $n = 1, 2, 3, \dots$.

---

*Pafnuty Lvovich Chebyshev (or Tschebyscheff), Russian mathematician (1821–1894).

†Carl Jacobi, German mathematician (1804–1851).

11. Apply Eq. (75) to evaluate $\int_{-\ell/2}^{+\ell/2} \psi_1(x) x \psi_2(x) \mathrm{d}x$ and $\int_{-\ell/2}^{+\ell/2} \psi_1(x) x^2 \psi_2(x) \mathrm{d}x$.
Ans. $16\ell/9\pi^2, 0$

12. Derive Eq. (81).

13. Show that $\xi^2 e^{-\frac{1}{2}\xi^2}$ is an asymptotic solution to Eq. (83) that leads to Hermite's equation.

14. Derive the recursion relation for the Hermite polynomials [Eq. (90)].

15. Derive Eqs. (97) and (99).

16. Develop Eqs. (101) and (102) and show that their substitution in Eq. (85) yields Eq. (99).

17. With the use of Eq. (111) prove Eq. (104).

18. Substitute Eq. (113) in Eq. (112) and derive Eq. (114).

19. Derive the recursion relation given by Eq. (118).

20. With the use of Eq. (121) derive Eqs. (122) to (124).

21. Develop the indicial equation for the associated Laguerre polynomials [Eq. (133)].

22. Derive the recursion relation [Eq. (136)] for the associated Laguerre polynomials.

23. Verify the relations between Eqs. (151) and (152).

24. Substitute $y = u\sqrt{x}$ in Eq. (144) to obtain Eq. (147).

6 Partial Differential Equations

Although the title of this chapter is general, it will be concerned only with the most important examples of partial differential equations of interest to physicists and chemists. Fortunately, the equations involved in virtually all of these applications can be solved by the very powerful method of separation of variables.

A partial differential equation is one with two or more independent variables. The separation of these variables, if it can be carried out, yields ordinary differential equations which can, in most cases, be solved by the various methods presented in Chapters 3 and 5. The general approach to this problem will now be illustrated by a number of examples that are fundamental in physics and chemistry.

6.1 THE VIBRATING STRING

Consider a flexible string of length ℓ that is stretched between two points by a constant tension τ. It will be assumed that the tension is sufficient so that the effect of gravity can be neglected. Furthermore, the string is uniform, with a density (mass per unit length) equal to ρ. The x axis is chosen along the direction of the string at rest and the displacement of the string is in the y direction.

6.1.1 The wave equation

Now consider the displacement of a segment of the string, Δs, as shown in Fig. 1. Its mass is equal to $\rho\Delta s$ and, according to Newton's second law of motion,

$$\tau_y(x+\Delta x) - \tau_y(x) = \rho\Delta s\frac{\partial^2 y}{\partial t^2}, \tag{1}$$

where $\partial^2 y/\partial t^2$ is its acceleration in the y direction. From Fig. 1 [see also Eq.(3-51)],

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 = (\Delta x)^2 + \left(\Delta x\frac{\partial y}{\partial x}\right)^2 \approx (\Delta x)^2, \tag{2}$$

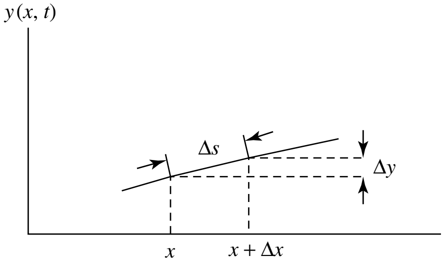


Fig. 1 Segment of a string.

as in the limit of small displacements, the slope approaches zero. Furthermore, in this limit $\tau_y = \tau\partial y/\partial x$ and $\tau_x \approx \tau$. Then Eq. (1) becomes

$$\tau\left[\left(\frac{\partial y}{\partial x}\right)_{x+\Delta x} - \left(\frac{\partial y}{\partial x}\right)_x\right] = \tau\Delta x\frac{\partial^2 y}{\partial x^2} = \rho\Delta x\frac{\partial^2 y}{\partial t^2}, \tag{3}$$

or,

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2}. \tag{4}$$

The quantity $c = \sqrt{\tau/\rho}$ is known as the phase velocity. It is the speed at which waves travel along the string. Clearly, the left-hand side of Eq. (4) represents the one-dimensional Laplacian operating on the dependent variable. This expression can be easily generalized to represent wave phenomena in two or more dimensions in space.

6.1.2 Separation of variables

In the application of the method of separation of variables to Eq. (4), it is assumed, without initial justification, that the dependent variable can be written as a product, *viz.*

$$y(x,t) = X(x)\vartheta(t). \tag{5}$$

Substitution of Eq. (5) into Eq. (4) yields

$$c^2\vartheta(t)\frac{\mathrm{d}^2X(x)}{\mathrm{d}x^2} = X(x)\frac{\mathrm{d}^2\vartheta(t)}{\mathrm{d}t^2}, \tag{6}$$

which, after division by $y(x,t) = X(x)\vartheta(t)$, becomes

$$\frac{c^2}{X(x)}\frac{\mathrm{d}^2X(x)}{\mathrm{d}x^2} = \frac{1}{\vartheta(t)}\frac{\mathrm{d}^2\vartheta(t)}{\mathrm{d}t^2}. \tag{7}$$

The left-hand side of Eq. (7) does not depend on the time t; it is only a function of the coordinate x. On the other hand, the right-hand side of this equation depends only on the time. As t and x are independent variables, each side of Eq. (7) must be equal to a constant. Furthermore, it must be the same constant, if Eq. (7) is to hold. This argument, which will be employed often in subsequent examples, is the basis of the method of the separation of variables. Clearly, the constant in question can be chosen at will. For convenience in this example, it will be set equal to $-\omega^2$.

The method illustrated above allows Eq. (7) to be decomposed into two ordinary differential equations, namely,

$$\frac{d^2\vartheta(t)}{dt^2} + \omega^2\vartheta(t) = 0 \tag{8}$$

and

$$\frac{d^2X(x)}{dx^2} + \frac{\omega^2}{c^2}X(x) = 0. \tag{9}$$

These two equations, which have the same form, have already been solved (see Section 5.2). One form of the general solution in each case is

$$\vartheta(t) = A\,sin\,\omega t + B\,cos\,\omega t \tag{10}$$

and

$$X(x) = C\,sin\left(\frac{\omega x}{c}\right) + D\,cos\left(\frac{\omega x}{c}\right), \tag{11}$$

respectively. The constants A and B appearing in Eq. (10) are of course the two arbitrary constants of integration arising from the general solution to the second-order differential equation for t. These constants can only be evaluated with the aid of the appropriate initial conditions.

6.1.3 Boundary conditions

In the present example it will be assumed that the string is fixed at each end, as is the case for musical instruments such as the violin and the guitar. Clearly, the string cannot vibrate at its ends; thus, $X(0) = 0$ and $X(\ell) = 0$ for a string of length ℓ (see Fig. 2). These conditions are imposed on the general solution in order to determine the constants of integration. From Eq. (11) it is evident that $X(0) = D = 0$. However, the remaining solution is $X(x) = C\,sin(\omega x/c)$, which must vanish at the other end of the string, where $x = \ell$. Clearly, C cannot be equated to zero, as the resulting solution is trivial; that is, $X(x) = 0$ for all values of x. However, as shown in Section 5.4.1, if the argument of sine is replaced by $n\pi x/\ell$, the condition $X(\ell) = 0$ will be fulfilled if n

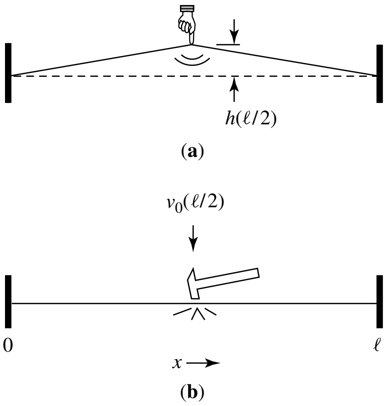


Fig. 2 String with fixed ends: (**a**) the plucked string; (**b**) the struck string.

is an integer. Thus, the solution for the spatial part of the problem is of the form

$$X(x) = C\,sin\left(\frac{n\pi x}{\ell}\right). \tag{12}$$

Equation (12) represents a standing wave in space with a wavelength λ determined by the condition $2\pi x/\lambda = n\pi x/\ell$ or $\lambda = 2\ell/n$. This result is, aside from notation, the same as that obtained for the quantum mechanical problem of the particle in a box [see Eq. (5-68) and Fig. 5-4a]. It should be noted that the vibrations of the string are quantized, although the problem is a classical one. The quantization arises in both classical and quantum-mechanical cases from the boundary conditions. The integers n, which arise naturally, determine the characteristic values – or often, eigenvalues (German: *Eigenwerte*). The corresponding functions, given in this case by Eq. (12), are the eigenfunctions. This subject will be developed in more detail in the following chapter.

A particular solution to Eq. (5) can now be written as

$$y_n(x,t) = X_n(x)\vartheta_n(t) = sin\left(\frac{n\pi x}{\ell}\right)\left[A_n{}'\,sin\left(\frac{n\pi ct}{\ell}\right) + B_n{}'\,cos\left(\frac{n\pi ct}{\ell}\right)\right], \tag{13}$$

where the coefficient C has been absorbed in the new constants, $A_n{}'$ and $B_n{}'$. The general solution is then given by

$$y(x,t) = \sum_{n=1}^{\infty} y_n(x,t) = \sum_{n=1}^{\infty} sin\left(\frac{n\pi x}{\ell}\right)\left[A_n{}'\,sin\left(\frac{n\pi ct}{\ell}\right) + B_n{}'\,cos\left(\frac{n\pi ct}{\ell}\right)\right]. \tag{14}$$

6.1.4 Initial conditions

The remaining arbitrary constants A'_n and B'_n are determined by the initial conditions. They depend on the manner in which the string is put into oscillation, as treated in Section 5.2.2 for the oscillation of the classical pendulum. There were two simple possibilities illustrated: (i) corresponds to a plucked string (as in the guitar) and leads to Eq. (5-36), while (ii) describes the action of a hammer in the piano, which strikes a string; the mathematical expression in this case is given by Eq. (5-38). These methods of exciting the vibration of the string are shown in Fig. 2.

With these ideas on mind, Eq. (14) can now be considered more generally. With the application of condition (i) at $t = 0$

$$y(x,0) = \sum_{n=1}^{\infty} B'_n \, sin\left(\frac{n\pi x}{\ell}\right). \tag{15}$$

This expression represents the expansion of an arbitrary function $y(x, 0)$ in a series of sines.* To determine the coefficient B'_n, multiply both sides of Eq. (15) by $sin(m\pi x/\ell)$ and integrate from $x = 0$ to $x = \ell$. Then,

$$\int_0^{\ell} y(x,0)\, sin\left(\frac{n\pi x}{\ell}\right) \mathrm{d}x = \sum_{n=1}^{\infty} B'_n \int_0^{\ell} sin\left(\frac{n\pi x}{\ell}\right) sin\left(\frac{m\pi x}{\ell}\right) \mathrm{d}x. \tag{16}$$

With m an integer the use of the relations

$$\int_0^{\ell} sin\left(\frac{n\pi x}{\ell}\right) sin\left(\frac{m\pi x}{\ell}\right) \mathrm{d}x = \begin{cases} 0, & \text{if } n \neq m \\ \ell/2, & \text{if } n = m \end{cases}, \tag{17}$$

yields the expression

$$B'_n = \frac{2}{\ell} \int_0^{\ell} y(x,0)\, sin\left(\frac{n\pi x}{\ell}\right) \mathrm{d}x. \tag{18}$$

The arbitrary constant B'_n is thus determined in the general solution given by Eq. (14). See problem 1.

As for the second arbitrary constant, recourse is made to condition (ii) above. Namely,

$$\dot{y}(x,0) = v_0(x) = \sum_{n=1}^{\infty} \frac{n\pi c}{\ell} A'_n \, sin\left(\frac{n\pi x}{\ell}\right), \tag{19}$$

---

*Equation (15) is an example of a Fourier series [Joseph Fourier, French mathematician (1768–1830)].

where $\dot{y}(x,0) = 0$ is the initial velocity. By a procedure similar to that employed above, the coefficients A'_n are found to be

$$A'_n = \frac{2}{n\pi c}\int_0^\ell v_0(x)\, sin\left(\frac{n\pi x}{\ell}\right) \mathrm{d}x \tag{20}$$

(problem 3). From Eqs. (18) and (20) it is evident that if the string has no initial velocity, the constants A'_n are equal to zero, while if the string has no initial displacement, the constants B'_n are equal to zero.

Each term in Eq. (14) represents a standing wave. For each value of n the frequency of vibration is given by

$$\nu_n^0 = \frac{nc}{2\ell} = \frac{n}{2\ell}\sqrt{\frac{\tau}{\rho}}. \tag{21}$$

The fundamental vibrational frequency is that with $n = 1$, while the frequencies of the harmonics or overtones are obtained with $n = 2, 3, 4 \ldots$. Specifically, $n = 2$ is called the "second harmonic" in electronics and the "first overtone" in musical acoustics. Both terms are employed, often erroneously, in the description of molecular vibrations (see Chapter 9).

As an example of the application of condition (i) above, consider the plucked string (see Fig. 2). The string is displaced at its midpoint by a distance h and released at $t = 0$. Thus, the initial conditions are

$$y(x,0) = \begin{cases} 2hx/\ell & 0 < x < \ell/2 \\ 2h(\ell - x)/\ell & \ell/2 < x < \ell \end{cases} \tag{22}$$

and $\dot{y}(x,0) = 0$. Substitution of Eq. (22) in Eq. (18) yields the relation

$$B'_n = \frac{2}{\pi}\left[\int_0^{\ell/2} \frac{2hx}{\ell}\, sin\left(\frac{n\pi x}{\ell}\right)\mathrm{d}x + \int_{\ell/2}^{\ell} \frac{2h}{\ell}(\ell - x)\, sin\left(\frac{n\pi x}{\ell}\right)\mathrm{d}x\right] \tag{23}$$

$$= \frac{8h}{\pi^2 n^2}\, sin\left(\frac{\pi n}{2}\right), \qquad \text{(with } n \text{ odd)} \tag{24}$$

for the integration constant. These results can be substituted into Eq. (14) to obtain

$$y(x,t) = \frac{8h}{\pi^2}\left[sin\left(\frac{\pi x}{\ell}\right) cos\left(\frac{\pi ct}{\ell}\right) - \frac{1}{9}\, sin\left(\frac{3\pi x}{\ell}\right) cos\left(\frac{3\pi ct}{\ell}\right) + \ldots\right], \tag{25}$$

which describes the vibration of the string after release from its initial position (problem 2). The first term represents the fundamental vibration, while the second corresponds to the second overtone ($n = 3$). The latter has an amplitude which is one ninth that of the fundamental, and thus a relative intensity of 1/81. The odd overtones (even harmonics), which have nodes in the center, are not excited because the string was plucked at that point (see Fig. 5-4a).

As musicians know, it is the relative intensities of the various members of the overtone series that determine the timbre or tone quality of sound. It is easy to distinguish the sound of a flute from that of the clarinet, although the listener may not know why. The sound of the flute has a relatively intense first overtone, while the boundary conditions imposed on the vibrating air column in the clarinet result in the suppression of all odd overtones. Such phenomena are of course much easier to visualize on a stringed instrument. Ask a violinist for a demonstration of the natural harmonics of a given string.

6.2 THE THREE-DIMENSIONAL HARMONIC OSCILLATOR

The classical harmonic oscillator in one dimension was illustrated in Section 5.2.2 by the simple pendulum. Hooke's law was employed in the form $f = -\kappa x$ where f is the force acting on the mass and κ is the force constant. The force can also be expressed as the negative gradient of a scalar potential function, $V(x) = \frac{1}{2}\kappa x^2$, for the problem in one dimension [Eq. (4-88)]. Similarly, the three-dimensional harmonic oscillator in Cartesian coordinates can be represented by the potential function

$$V(x, y, z) = \tfrac{1}{2}\kappa_x x^2 + \tfrac{1}{2}\kappa_y y^2 + \tfrac{1}{2}\kappa_z z^2, \tag{26}$$

where the force constants κ_x, κ_y and κ_z define Hooke's law in the corresponding directions.

6.2.1 Quantum-mechanical applications

In the analogous quantum-mechanical problem the kinetic energy of the system is represented by the operator $-(\hbar^2/2m)\nabla^2$, as developed in the following chapter. Its one-dimensional analog was already employed in Eq. (5-64). Thus, the Schrödinger equation for the three-dimensional harmonic oscillator is given by

$$-\frac{\hbar^2}{2m}\nabla^2\psi(x, y, z) + (\tfrac{1}{2}\kappa_x x^2 + \tfrac{1}{2}\kappa_y y^2 + \tfrac{1}{2}\kappa_z z^2)\psi(x, y, z) = \varepsilon\psi(x, y, z). \tag{27}$$

With the Laplacian written in Cartesian coordinates, Eq. (27) becomes

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+\frac{1}{2}\kappa_x x^2\right)\psi(x,y,z)+\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial y^2}+\frac{1}{2}\kappa_y y^2\right)\psi(x,y,z)$$
$$+\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2}+\frac{1}{2}\kappa_z z^2\right)\psi(x,y,z)=\varepsilon\psi(x,y,z). \tag{28}$$

The separation of variables is accomplished by substituting

$$\psi(x,y,z)=X(x)Y(y)Z(z) \tag{29}$$

and dividing by the same expression. The result is

$$\frac{1}{X(x)}\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+\frac{1}{2}\kappa_x x^2\right)X(x)+\frac{1}{Y(y)}\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial y^2}+\frac{1}{2}\kappa_y y^2\right)Y(y)$$
$$+\frac{1}{Z(z)}\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2}+\frac{1}{2}\kappa_z z^2\right)Z(z)=\varepsilon. \tag{30}$$

Each term on the left-hand side of Eq. (30) is a function of only a single independent variable. Each term is, therefore, equal to a constant, such that $\varepsilon_x+\varepsilon_y+\varepsilon_z=\varepsilon$. The first term is identified as

$$\frac{1}{X(x)}\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+\frac{1}{2}\kappa_x x^2\right)X(x)=\varepsilon_x, \tag{31}$$

an ordinary, second-order differential equation. Analogous relations for the terms in y and z are easily obtained.

To put Eq. (31) into a recognizable form, it is convenient to change the independent variable by substituting $\xi=2\pi x\sqrt{\nu_x^0 m/h}$, where $\nu_x^0=\sqrt{\kappa_x/m}/2\pi$ is the classical frequency of oscillation in the x direction (see Section 5.2.2). Then with $\sigma=2\varepsilon_x/h\nu_x^0$, Eq. (31) becomes

$$\frac{\partial^2 X(\xi)}{\partial\xi^2}+(\sigma-\xi^2)X(\xi)=0, \tag{32}$$

which, aside from notation, is the same as Eq. (5-82). Its solution can then be expressed in terms of the Hermite polynomials, with the energy given by Eq. (5-92) in the form

$$\varepsilon_x=h\nu_x^0(v_x+\tfrac{1}{2}), \tag{33}$$

where $v_x=0,1,2,\ldots$, the vibrational quantum number in the x direction. Clearly, the same procedure can be applied to the equations for $Y(y)$ and $Z(z)$, with similar results. The total energy for the three-dimensional oscillator is then given by

$$\varepsilon = h\nu_x^0(v_x + \tfrac{1}{2}) + h\nu_y^0(v_y + \tfrac{1}{2}) + h\nu_z^0(v_z + \tfrac{1}{2}). \tag{34}$$

6.2.2 Degeneracy

An example of the application of Eq. (34) is shown in Fig. 3, where the energy has been calculated for various values of the three (independent) quantum numbers v_x, v_y and v_z. The classical vibrational frequencies were chosen so that $v_x + v_y + v_z =$ constant. Thus, the minimum energy, that of the ground state (0, 0, 0), is obtained when all three quantum numbers are equal to zero and is the same for all combinations shown in Fig. 3. In the first case, in which $\nu_x^0 \neq \nu_y^0 \neq \nu_z^0$ the frequencies were arbitrarily chosen in the proportion $v_x : v_y : v_z = 1.0 : 1.1 : 1.2$. The resulting energy levels are shown in the figure. It should be noted that on several cases, *e.g.* the levels 0, 2, 0; 1, 0, 1, two different combinations of the three quantum numbers yield the same value for the energy. These levels are said to be degenerate, that is, two different wavefunctions yield exactly the same energy. In this case the degeneracy is

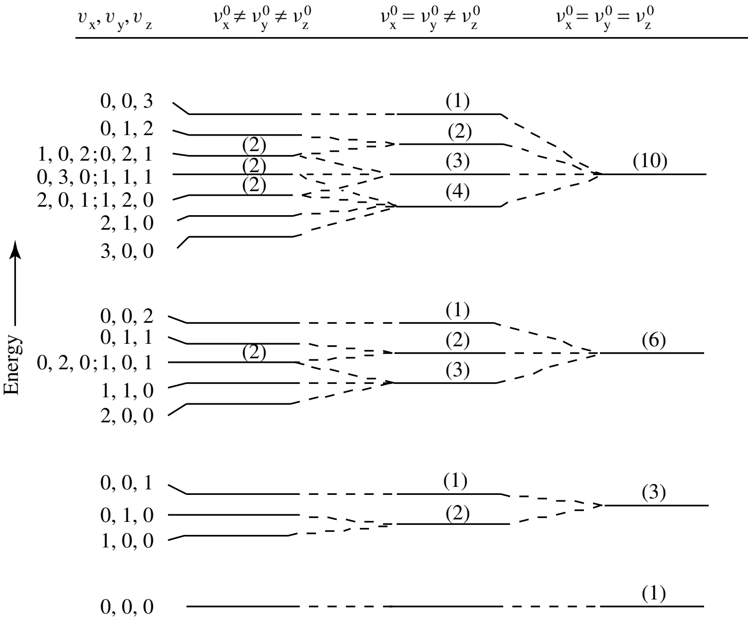


Fig. 3 Energy levels of the three-dimensional harmonic oscillator. The degree of degeneracy of each level is shown in parenthesis.

due to the particular values of the vibrational frequencies, so it is called an "accidental" degeneracy. Since two different wavefunctions contribute to the degenerate pair, the level is doubly degenerate, as indicated by (2) in Fig. 3.

A second combination of the vibrational frequencies is that in which $\nu_x^0 = \nu_y^0 \neq \nu_z^0$. The calculated energy levels are shown in the second column of Fig. 3. Here, the proportion of the vibrational frequencies has been chosen to be $v_x : v_y : v_z = 1.05 : 1.05 : 1.2$. The system now has a natural symmetry, since the two directions x and y are equivalent. The result is an increase in the degeneracy of the vibrational levels, an important consequence of the symmetry of the problem.

Finally, the combination for which $\nu_x^0 = \nu_y^0 = \nu_z^0$ corresponds to an isotropic potential, one in which the three spatial directions are equivalent. The resulting energy levels are shown in the last column of Fig. 3, where the vibrational frequencies have been chosen in the proportions $v_x : v_y : v_z = 1.1 : 1.1 : 1.1$. The degree of degeneracy for each energy level is shown in parentheses. Clearly, the increased degeneracy of the system is the result of the increased symmetry. This problem will be analyzed with the aid of the theory of groups in Chapter 8.

The energy of the isotropic harmonic oscillator in three dimensions can be written as

$$\varepsilon = h\nu^0(v_x + v_y + v_z + \tfrac{3}{2}) = h\nu^0(v + \tfrac{3}{2}) \tag{35}$$

where $v = v_x + v_y + v_z$. Thus, for a given value of v, v_x can take the values $v_x = 0, 1, 2, \ldots v$, or $v + 1$ different values. Then, $v_y = 0, 1, 2, \ldots v - v_x$, that is, $v + 1 - v_x$ values, leaving only one possibility for v_z, namely, $v_z = v - v_x - v_y$. Hence the total number of combinations of the three quantum numbers for a given value of v is given by

$$\begin{aligned}\sum_{v_x=0}^{v}(v + 1 - v_x) &= (v+1)\sum_{v_x=0}^{v} 1 - \sum_{v_x=0}^{v} v_x = (v+1)(v+1) - \frac{v(v+1)}{2} \\ &= (v+1)\left(\frac{v}{2} + 1\right).\end{aligned} \tag{36}$$

This expression* has been used to calculate the degeneracies shown in parentheses in the last column of Fig. 3.

---

*Note the general relations $\sum_{k=0}^{n} 1 = n + 1$ and $\sum_{k=0}^{n} k = n(n+1)/2$ that were used in deriving Eq. (36).

6.3 THE TWO-BODY PROBLEM

6.3.1 Classical mechanics

Consider a system composed of two particles of masses m_1 and m_2 in three dimensions. Three coordinates are necessary to specify the position of each particle. In a Cartesian coordinate system the total energy can be written as

$$\varepsilon = \frac{m_1}{2}(\dot{x}_1^2 + \dot{y}_1^2 + \dot{z}_1^2) + \frac{m_2}{2}(\dot{x}_2^2 + \dot{y}_2^2 + \dot{z}_2^2) + V(x_1, y_1, z_1, x_2, y_2, z_2), \quad (37)$$

where the first two terms in Eq. (37) represent the classical kinetic energy of the system and the third the potential energy. The positions of the particles with respect to the origin fixed in space (O in Fig. 4) are specified by the vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, whose components are x_1, y_1, z_1 and x_2, y_2, z_2, respectively. The position of the center of mass (cm) is defined by the relation $|\boldsymbol{a}_1|m_1 = |\boldsymbol{a}_2|m_2$. The vector $\boldsymbol{r} = \boldsymbol{a}_1 + \boldsymbol{a}_2$ represents the separation between the two particles, which have here been assumed to be spherical. Thus, the relations $\boldsymbol{a}_1 = [m_2/(m_1 + m_2)]\boldsymbol{r}$ and $\boldsymbol{a}_2 = [m_1/(m_1 + m_2)]\boldsymbol{r}$ define the center of mass of the two-particle system. By inspection of the two triangles in Fig. 4 the following vector relations are easily established: $\boldsymbol{R} = \boldsymbol{r}_1 + \boldsymbol{a}_1$ and $\boldsymbol{R} = \boldsymbol{r}_2 - \boldsymbol{a}_2$. Substitution for $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ leads to the expressions

$$\boldsymbol{r}_1 = \boldsymbol{R} - \frac{m_2}{m_1 + m_2}\boldsymbol{r} \quad (38)$$

and

$$\boldsymbol{r}_2 = \boldsymbol{R} + \frac{m_1}{m_1 + m_2}\boldsymbol{r}. \quad (39)$$

In terms of components, Eqs. (38) and (39) correspond to six relations such as

$$x_1 = X - \frac{\mu}{m_1}x \quad (40)$$

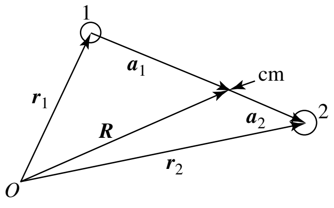


Fig. 4 The two-body problem.

and

$$x_2 = X + \frac{\mu}{m_2}x, \tag{41}$$

where X, Y, Z are the Cartesian components of the vector $\boldsymbol{R}$ which specify the position of the center of mass. Analogous relations are easily written for the other components. The reduced mass, defined by $\mu = m_1m_2/(m_1 + m_2)$, has been introduced in Eqs. (40) and (41). From Eq. (37) the kinetic energy is given by

$$T = \frac{m_1}{2}(\dot{x}_1^2 + \dot{y}_1^2 + \dot{z}_1^2) + \frac{m_2}{2}(\dot{x}_2^2 + \dot{y}_2^2 + \dot{z}_2^2)$$

or

$$T = \frac{\mu}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{M}{2}(\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2), \tag{42}$$

after substitution of the time derivatives of Eqs. (40) and (41) and writing the total mass as $M = m_1 + m_2$. The kinetic energy can equally be expressed in terms of the momenta. Thus, the components of the conjugate momenta are of the form

$$p_x \equiv \frac{\partial T}{\partial \dot{x}} = \mu\dot{x}, \tag{43}$$

and similarly for the five others. Their substitution in Eq. (42) results in an expression for the kinetic energy as a function of the momenta. This step is essential before the transformation of the classical formulation into the quantum-mechanical one. The result in this case is given by

$$T = \frac{\boldsymbol{p}^2}{2\mu} + \frac{\boldsymbol{P}^2}{2M} \tag{44}$$

(problem 6).

The classical kinetic energy of the system has now been separated into the effect of displacement of the center of mass of the system, with momentum $\boldsymbol{P}$ and that of the relative movement of the two particles, with momentum $\boldsymbol{p}$. In the absence of external forces, the interaction of the two (spherical) particles is only a function of their separation, $\boldsymbol{r}$. That is, the potential function appearing in Eq. (37) depends only on the "internal" coordinates x, y, z.

6.3.2 Quantum mechanics

In the quantum mechanical applications of the two-body problem, the classical energy of the system becomes the Hamiltonian operator.* The conversion

---

*William Rowan Hamilton, Irish mathematician and astronomer (1805–1865).

is accomplished by replacing each momentum vector by the corresponding operator, as shown in the following chapter, *viz.*

$$\boldsymbol{p} \to \frac{\hbar}{i}\nabla_{x,y,z} \tag{45}$$

and

$$\boldsymbol{P} \to \frac{\hbar}{i}\nabla_{X,Y,Z}. \tag{46}$$

Substitution in Eq. (44) yields the Hamiltonian for this problem,

$$\hat{H} = \hat{T} + V(x, y, z) = -\frac{\hbar^2}{2\mu}\nabla^2_{x,y,z} - \frac{\hbar^2}{2M}\nabla^2_{X,Y,Z} + V(x, y, z). \tag{47}$$

The Schrödinger equation, with ε the total energy, is then

$$\hat{H}\varsigma(x, y, z, X, Y, Z) = \varepsilon\varsigma(x, y, z, X, Y, Z). \tag{48}$$

It can be separated by the substitution

$$\varsigma(x, y, z, X, Y, Z) = \psi(x, y, z)\vartheta(X, Y, Z), \tag{49}$$

followed by division by the same function. The result can be written in the form

$$\begin{aligned} &-\frac{1}{\psi(x, y, z)}\frac{\hbar^2}{2\mu}\nabla^2_{x,y,z}\psi(x, y, z) + V(x, y, z) \\ &-\frac{\hbar^2}{\vartheta(X, Y, Z)2M}\nabla^2_{X,Y,Z}\vartheta(X, Y, Z) = \varepsilon \end{aligned} \tag{50}$$

The first two terms on the left-hand side of Eq. (50) are functions only of the internal coordinates, while the third term depends only on the external coordinates X, Y, Z. Therefore, each must be equal to a constant, such that their sum is equal to ε. Thus, if

$$-\frac{1}{\psi(x, y, z)}\frac{\hbar^2}{2\mu}\nabla^2_{x,y,z}\psi(x, y, z) + V(x, y, z) = \varepsilon_{int}, \tag{51}$$

$$-\frac{\hbar^2}{\vartheta(X, Y, Z)2M}\nabla^2_{X,Y,Z}\vartheta(X, Y, Z) = \varepsilon - \varepsilon_{int} = \varepsilon_{ext}, \tag{52}$$

and the separation of the internal and external coordinates has been accomplished. After multiplication of Eq. (52) by $\vartheta(X, Y, Z)$ it can be recognized as the Schrödinger equation for a free particle of mass $M = m_1 + m_2$ and energy equal to ε_{ext}. This energy is not quantized unless boundary conditions are applied. The problem can be further separated into equations for the particle in each Cartesian direction. It should be noted that the separation of variables to yield Eqs. (51) and (52) is only possible because the potential function for the free particle is independent of the external coordinates.

Multiplication of Eq. (51) by $\psi(x, y, z)$ yields the Schrödinger equation for the relative movement of the two particles. However, the Cartesian coordinates employed are not "natural" for this problem. In particular, if, as has been assumed, the particles are spherical, the interaction potential depends only on their distance of separation, r. The problem then reduces to the movement of a hypothetical particle of mass μ in the central field of a potential $V(r)$. The various applications of this result depend specifically on the form of this potential function.

6.4 CENTRAL FORCES

6.4.1 Spherical coordinates

With Eq. (51) written in the form

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2_{x,y,z} + V(r)\right]\psi(x, y, z) = \varepsilon_{int}\psi(x, y, z), \tag{53}$$

it is now necessary to convert the Laplacian operator into spherical polar coordinates, which correspond to the symmetry of the potential function. This operation can be carried out by direct substitution of the relations

$$\begin{aligned} x &= r\, sin\,\theta\, cos\,\varphi \\ y &= r\, sin\,\theta\, sin\,\varphi \\ z &= r\, cos\,\theta, \end{aligned} \tag{54}$$

where the new coordinates r, θ, φ are defined in Fig. 5. The direct change of variables is given in Appendix V. However, the method developed in Chapter 4 is much easier. With the use of Eq. (54) the appropriate scale factors calculated from Eq. (4-73) are: $h_r = 1$, $h_\theta = r$ and $h_\varphi = r\,sin\,\theta$. Substitution of these quantities in Eq. (4-100) leads directly to the result

$$\nabla^2_{r,\theta,\varphi} = \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{sin\,\theta}\frac{\partial}{\partial\theta}\left(sin\,\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] \tag{55}$$

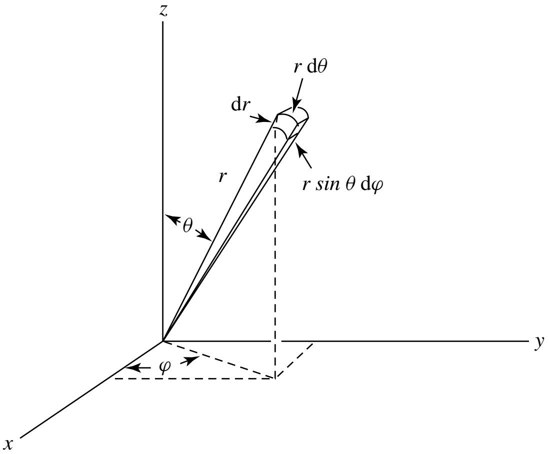


Fig. 5 Spherical coordinates.

for the Laplacian operator and

$$\mathrm{d}\tau = h_r h_\theta h_\varphi \mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi = r^2\,sin\,\theta\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi \tag{56}$$

for the volume element [see Eq. (4-96)].

Equation (52) can now be rewritten in spherical coordinates as

$$\left\{-\frac{\hbar^2}{2\mu r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{1}{sin\,\theta}\frac{\partial}{\partial\theta}\left(sin\,\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{sin^2\,\theta}\frac{\partial^2}{\partial\varphi^2}\right]+V(r)\right\}$$
$$\times\,\psi(r,\theta,\varphi)=\varepsilon_{int}\psi(r,\theta,\varphi). \tag{57}$$

This form of Schrödinger's equation can be separated with the use of the substitution

$$\psi(r,\theta,\varphi)=R(r)Y(\theta,\varphi). \tag{58}$$

The result is the partial differential equation for $Y(\theta,\varphi)$,

$$\left[\frac{1}{sin\,\theta}\frac{\partial}{\partial\theta}\left(sin\,\theta\frac{\partial Y(\theta,\varphi)}{\partial\theta}\right)+\frac{1}{sin^2\theta}\frac{\partial^2 Y(\theta,\varphi)}{\partial\varphi^2}\right]=-\beta Y(\theta,\varphi), \tag{59}$$

and the ordinary differential equation for $R(r)$,

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R(r)}{\mathrm{d}r}\right)+\frac{2\mu r^2}{\hbar^2}[\varepsilon_{int}-V(r)]R(r)=\beta R(r). \tag{60}$$

6.4.2 Spherical harmonics

Equation (59) can be further separated by substituting $Y(\theta, \varphi) = \Theta(\theta)\Phi(\varphi)$ and multiplying by $sin^2 \theta$. After division by $\Theta(\theta)\Phi(\varphi)$, the result is

$$\frac{sin\,\theta}{\Theta(\theta)}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(sin\,\theta\frac{\mathrm{d}\Theta(\theta)}{\mathrm{d}\theta}\right) + \frac{1}{\Phi(\varphi)}\frac{\mathrm{d}^2\Phi(\varphi)}{\mathrm{d}\varphi^2} + \beta\,sin^2\theta = 0. \tag{61}$$

Clearly, the second term on the left-hand side is equal to a constant, say $-m^2$. The resulting equation for $\Phi(\varphi)$,

$$\frac{\mathrm{d}^2\Phi(\varphi)}{\mathrm{d}\varphi^2} + m^2\Phi(\varphi) = 0, \tag{62}$$

is of the same form as Eq. (5-29). A particular solution is, therefore, an exponential function of an imaginary argument, *viz.*

$$\Phi(\varphi) = e^{im\varphi}. \tag{63}$$

The coordinate φ is cyclic in the sense that on physical grounds the exponential must have the same value when φ is advanced by 2π. That is, the condition $\Phi(\varphi) = \Phi(\varphi + 2\pi)$ must be fulfilled for the function to be single-valued. Then,

$$\Phi(\varphi) = e^{im\varphi} = e^{im(\varphi+2\pi)} = e^{im\varphi}e^{2\pi im}, \tag{64}$$

which implies $e^{2\pi im} = 1$, and restricts m to the values $0, \pm1, \pm2, \ldots$. The normalization of this function is accomplished with the use of the factor $\mathcal{N}_\varphi$ such that

$$\int_0^{2\pi} \Phi^\star\Phi\mathrm{d}\varphi = \mathcal{N}_\varphi^2\int_0^{2\pi} e^{-im\varphi}e^{im\varphi}\mathrm{d}\varphi = 2\pi\mathcal{N}_\varphi^2 = 1. \tag{65}$$

The functions are then given by

$$\Phi(\varphi) = \frac{1}{\sqrt{2\pi}}e^{im\varphi}, \tag{66}$$

with $m = 0, \pm1, \pm2, \ldots$, as before.

The substitution of $-m^2$ for the second term on the left-hand side of Eq. (60) yields the equation for $\Theta(\theta)$,

$$\frac{1}{sin\,\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(sin\,\theta\frac{\mathrm{d}\Theta(\theta)}{\mathrm{d}\theta}\right) - \frac{m^2\Theta(\theta)}{sin^2\theta} + \beta\Theta(\theta) = 0. \tag{67}$$

It is convenient to make the substitutions $z = cos\,\theta$ and $\Theta(\theta) = P(z)$. The result,

$$(1 - z^2)\frac{\mathrm{d}^2 P(z)}{\mathrm{d}z^2} - 2z\frac{\mathrm{d}P(z)}{\mathrm{d}z} + \left(\beta - \frac{m^2}{1 - z^2}\right) P(z) = 0, \tag{68}$$

is identical to Eq. (5-112). Its solution is expressed in terms of the associated Legendre polynomials, which when normalized, are the functions $\Theta_{\ell,m}(\theta)$, with $\ell = 0, 1, 2, \ldots$ and $m = 0, \pm 1, \pm 2, \ldots, \pm\ell$. Furthermore, the separation constant can be identified as $\beta = \ell(\ell + 1)$, as given by Eq. (5-119). The products

$$\Phi_m(\varphi)\Theta_{\ell,m}(\theta) = Y_\ell^m(\theta, \varphi) \tag{69}$$

are the spherical harmonics [see Eq. (5-129)]. These functions are solutions for the angular dependence of the wavefunction for all central force problems. In real form they are often referred to as atomic orbitals (see Appendix III).

The radial part of the wavefunction depends on the potential function that describes the interaction of the two particles. Several examples which are important in chemistry and physics will now be summarized.

6.5 THE DIATOMIC MOLECULE

Within the framework of the Born–Oppenheimer approximation*, a diatomic molecule consists of two nuclei that are more-or-less attached by the surrounding electron cloud. Often the specific form of the resulting potential function is not known. However, if a chemical bond is formed between the two nuclei, the potential function displays a minimum at a distance that corresponds to the equilibrium bond length. Furthermore, the energy necessary to break the chemical bond, the dissociation energy, is often evaluated by spectroscopic measurements. It can be concluded, then, that the potential function has the general form shown in Fig. 6. A simple derivation of the Born–Oppenheimer approximation is presented in Section 12.1.

In this application Eq. (60) becomes

$$\left[-\frac{\hbar^2}{2\mu r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R(r)}{\mathrm{d}r}\right) + \frac{\hbar^2 J(J+1)}{2\mu r^2} + V(r)\right] R(r) = \varepsilon_{int} R(r), \tag{70}$$

where μ is now the reduced mass of the two nuclei and, by tradition, the quantum number ℓ has been replaced by the letter J.

---

*Max Born, British physicist (1882–1970); Julius Robert Oppenheimer, American physicist (1904–1967).

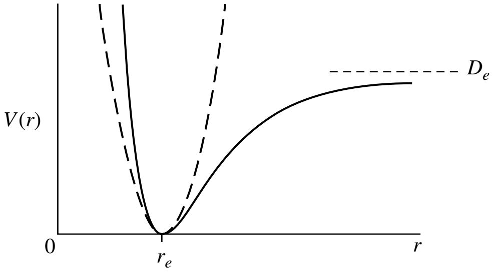


Fig. 6 Potential functions for a diatomic molecule. The dashed curve represents the harmonic approximation.

6.5.1 The rigid rotator

In the simplest case the interatomic distance is held constant, *e.g.* $r = r_e$ and the potential energy is set equal to zero at this point. Furthermore, as $r = r_e$ is constant, the first term on the left-hand side of Eq. (70) vanishes. These conditions describe the rigid rotator, for which the energy is given by

$$\varepsilon_{rr} = \frac{\hbar^2 J(J+1)}{2\mu r_e^2}, \tag{71}$$

with $J = 0, 1, 2, \ldots$, as before. The quantity $I_e = \mu r_e^2$ is the moment of inertia of the rigid diatomic molecule.

6.5.2 The vibrating rotator

Returning now to the general expression for $R(r)$ [Eq. (70)], it is convenient to change the dependent variable by substituting $R(r) = (1/r)S(r)$. The result is

$$\frac{\mathrm{d}^2 S(r)}{\mathrm{d}r^2} + \frac{2\mu}{\hbar^2}\left[-\frac{\hbar^2 J(J+1)}{2\mu r^2} + \varepsilon_{int} - V(r)\right]S(r) = 0. \tag{72}$$

As the molecule executes small-amplitude vibrations with respect to the equilibrium internuclear distance, it is appropriate to develop the potential function in a Taylor series about that position. Thus,

$$V(r) = V_e + \left.\frac{\mathrm{d}V}{\mathrm{d}r}\right|_e (r - r_e) + \frac{1}{2}\left.\frac{\mathrm{d}^2 V}{\mathrm{d}r^2}\right|_e (r - r_e)^2 + \cdots \ . \tag{73}$$

The potential energy can be set equal to zero at the equilibrium position; then, $V_e = 0$. Furthermore, at equilibrium the potential is minimal,

$$\left.\frac{\mathrm{d}V}{\mathrm{d}r}\right|_e = 0.$$

And, in the harmonic approximation cubic and higher terms are neglected, so that Eq. (73) becomes

$$V(r) = \frac{1}{2}\kappa(r - r_e)^2, \tag{74}$$

the harmonic potential function shown in Fig. 6. The force constant is defined by

$$\kappa \equiv \left.\frac{\mathrm{d}^2 V}{\mathrm{d}r^2}\right|_e,$$

the curvature of the potential function evaluated at the equilibrium position. Higher terms in Eq. (73) contribute to the anharmonicity of the vibration. This question will be discussed in Chapter 9.

In the rigid-rotator, harmonic-oscillator approximation Eq. (72) becomes

$$\frac{\mathrm{d}^2 S(r)}{\mathrm{d}r^2} + \frac{2\mu}{\hbar^2}\left[-\frac{\hbar^2 J(J+1)}{2\mu r^2} + \varepsilon_{int} - V(r)\right] S(r) = 0, \tag{75}$$

where $\varepsilon = \varepsilon_{vib} + \varepsilon_{rr}$ and $x = r - r_e$. Equation (75) can be put into the form of Eq. (32) by analogous substitutions. Thus, $\varepsilon_{vib} = h\nu^0(v + \frac{1}{2})$, with $\nu^0 = \sqrt{\kappa/\mu}/2\pi$ and $v = 0, 1, 2, \ldots$, as before. The result yields an expression which is the sum of the energy of a harmonic oscillator and that of a rotating molecule which does not oscillate! In spite of this apparent contradiction, the result is the starting point for the interpretation of the rotation–vibration spectrum of a diatomic molecule, as observed, for example, in the mid-infrared spectral region.

6.5.3 Centrifugal forces

A simple improvement on this model can be made by remarking that the first term in brackets in Eq. (72) contains the factor $1/r^2$. As the amplitude of the vibration is small, a binomial series development can be made (see Section 2.10), namely,

$$\frac{1}{r^2} = \frac{1}{(r_e + x)^2} = \frac{1}{r_e^2}\left(1 - \frac{2x}{r_e} + \frac{3x^2}{r_e^2} - \cdots\right), \tag{76}$$

where $r = r_e + x$. Clearly, the rigid-rotator approximation employed above corresponds to $x = 0$ in Eq. (76). If, however, the linear term in x is retained, the effect of the centrifugal force can be estimated.

Reconsideration of Eq. (72) with

$$\frac{1}{r^2} = \frac{1}{r_e^2}\left(1 - \frac{2x}{r_e}\right),$$

suggests a suitable substitution by translation of the x axis, namely $x = \zeta + a$, that will immediately simplify the problem. The constant a is chosen by setting the resulting linear terms in ζ equal to zero. The result is

$$a = -\frac{\hbar^2 J(J+1)}{\kappa\mu r_e^3}. \tag{77}$$

With $\varepsilon = \varepsilon_{vib} - \varepsilon_{rr}$, as before, Eq. (75) becomes

$$\frac{\mathrm{d}^2 S(\varsigma)}{\mathrm{d}\varsigma^2} + \frac{2\mu}{\hbar^2}\left[\varepsilon_{vib} - \frac{\hbar^4 J^2(J+1)^2}{2\kappa\mu^2 r_e^6} - \frac{1}{2}\kappa\varsigma^2\right] S(\varsigma) = 0. \tag{78}$$

This equation can be identified as that of the harmonic oscillator, with a supplementary constant term inside the brackets. The energy of the rotating, vibrating molecule is then given by

$$\varepsilon = h\nu^0\left(v + \frac{1}{2}\right) + \frac{\hbar^2 J(J+1)}{2\mu r_e^2} - \frac{\hbar^4 J^2(J+1)^2}{2\kappa\mu^2 r_e^6}. \tag{79}$$

The first term, with $v = 0, 1, 2, \ldots$, is the energy of the harmonic oscillator. The second, with $J = 0, 1, 2, \ldots$, is that of the rigid rotator, while the last term expresses the nonrigidity of the rotating molecule. Classically speaking, as the molecule turns more rapidly, the bond length increases due to centrifugal force and, thus, the energy decreases – as expressed by the negative sign in Eq. (79).

The energy of the diatomic molecule, as given by Eq. (79) does not take into account the anharmonicity of the vibration. The effect of the cubic and quartic terms in Eq. (73) can be evaluated by application of the theory of perturbation (see Chapter 12).

6.6 THE HYDROGEN ATOM

The representation of the angular part of the two-body problem in spherical harmonics, as developed in Section 6.4, is applicable to any system composed of two spherical particles in free space. For the hydrogen atom, composed of a proton and an electron, the reduced mass is equal to

$$\mu = \frac{m_e m_p}{m_e + m_p} \approx m_e, \tag{80}$$

where the approximation in which μ is replaced by the mass electronic m_e is satisfactory in most chemical applications (see problem 9).

6.6.1 Energy

The interaction between the two particles in this system is described by Coulomb's law,* in which the force is proportional to the inverse-square of the distance between the particles and $-e^2$ is the product of the charges on the electron and the proton. The corresponding potential function is then of the form

$$V(r) = -\frac{e^2}{4\pi\varepsilon_0 r}, \tag{81}$$

as, $\boldsymbol{f} = -\nabla V(r)$. The constant $4\pi\varepsilon_0$ in the denominator of Eq. (81) arises if international units are employed. With this potential function, as shown in Fig. 7, the radial equation [Eq. (60)] can be written for the hydrogen atom as

$$\left[-\frac{\hbar^2}{2\mu r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}}{\mathrm{d}r}\right) + \frac{\hbar^2\ell(\ell+1)}{2\mu r^2} - \frac{e^2}{4\pi\varepsilon_0 r}\right]R(r) = \varepsilon R(r), \tag{82}$$

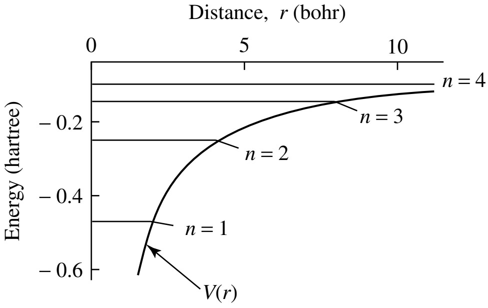


Fig. 7 The potential function $V(r)$ and the energy levels for the hydrogen atom.

---

*Charles de Coulomb, French physicist (1736–1801).

where β has been replaced by $\ell(\ell+1)$. It is now convenient to make the substitutions

$$\alpha^2 = -\frac{2\mu\varepsilon_{int}}{\hbar^2} \text{ and } \gamma = \frac{\mu e^2}{\alpha\hbar^2}.$$

Then, with the change in the independent variable, $\rho = 2\alpha r$, Eq. (82) becomes

$$\frac{\mathrm{d}^2R(\rho)}{\mathrm{d}\rho^2} + \frac{2}{\rho}\frac{\mathrm{d}R(\rho)}{\mathrm{d}\rho} + \left[\frac{\gamma}{\rho} - \frac{\ell(\ell+1)}{\rho^2} - \frac{1}{4}\right]R(\rho) = 0, \tag{83}$$

which is identical to Eq. (5-130). The solutions are then given by the Laguerre polynomials, as summarized in Section 5.5.3. There it was shown that because of the boundary conditions, γ is equal to a positive integer which was identified as the principal quantum number, n. Then, from the substitutions made above, the energy of the hydrogen atom is given by

$$\varepsilon_n = -\left(\frac{\mu e^4}{32\pi^2\varepsilon_0^2\hbar^2}\right)\frac{1}{n^2}, \tag{84}$$

with of course $n = 1, 2, 3, \ldots$. Some of the values of the energy are indicated in Fig. 7, where it is seen that with $n = 1$ the ground-state energy is equal to -13.6 eV or -0.5 hartree.

The spectrum of atomic hydrogen, as observed in absorption or emission, arises from transitions between the various possible states. In emission, a spectral line results from a transition such as $n_2 \to n_1$ and the application of Eq. (84) leads to the expression

$$\frac{\Delta\varepsilon}{hc} = R_H\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right), \tag{85}$$

where R_H is known as the Rydberg constant* (see problem 10).

6.6.2 Wavefunctions and the probability density

The radial parts of the wavefunctions for the hydrogen atom can be constructed from the general form of the associated Laguerre polynomials, as developed in Section 5.5.3. However, in applications in physics and chemistry it is often the probability density that is more important (see Section 5.4.1). This quantity in this case represents the probability of finding the electron in the appropriate three-dimensional volume element.

---

*Johannes Robert Rydberg, Swedish physicist (1854–1919).

As a simple example, consider the hydrogen atom in its ground state, $n = 1$. The radial part of the wavefunction is given by

$$R_{1,0}(r) = 2a_0^{-3/2} e^{-r/a_0}, \tag{86}$$

where

$$a_0 = \frac{4\pi\varepsilon_0\hbar^2}{\mu e^2} = 0.53 \text{ Å} = 1 \text{ bohr}$$

is the radius of the first Bohr orbit.* Although this quantity has no direct significance in modern quantum theory, it serves as a useful measure of distance on the atomic scale (see Appendix II). A number of the radial functions are given in Appendix IV. For the ground state the probability density is then of the form

$$P_{1,0}(r) = [R_{1,0}(r)]^2 r^2 = \frac{4r^2}{a_0^3} \mathrm{e}^{-2r/a_0}. \tag{87}$$

It is plotted in Fig. 8.

It is of interest to determine the position of the maximum of the function $P_{1,0}(r)$, as this distance describes the effective radius of the hydrogen atom in its ground state. The derivative of Eq. (87) is equal to

$$\frac{\mathrm{d}P_{1,0}(r)}{\mathrm{d}r} = \frac{8r}{a_0^3}\left(1 - \frac{r}{a_0}\right) \mathrm{e}^{-2r/a_0}. \tag{88}$$

This function is equal to zero at the origin, at infinity and, of course at $r = a_0 = 0.53$ Å, the position of the maximum. It is perhaps surprising that the

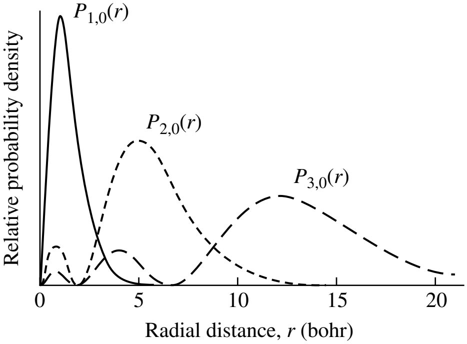


Fig. 8 Radial probability density for the hydrogen atom in *n*s states.

---

*Niels Bohr, Danish physicist (1885–1962).

most probable radial distance of the electron corresponds exactly to the radius of the first orbit in the old quantum theory. However, it must be emphasized that the probability distribution is now spherical, with a diameter of $2a_0$, or approximately one Ångström* (10^{-10} m) for the atom in its ground state.

6.7 BINARY COLLISIONS

6.7.1 Conservation of angular momentum

The interaction of two particles was analyzed classically in Section 6.3.1. The resulting expression for the relative momentum can be written in vector notation as

$$\boldsymbol{p} = \mu\dot{\boldsymbol{r}}. \tag{89}$$

Then, following Newton,

$$\dot{\boldsymbol{p}} = \mu\ddot{\boldsymbol{r}} = \boldsymbol{f} \tag{90}$$

or

$$\ddot{\boldsymbol{r}} = \frac{\boldsymbol{f}}{\mu}. \tag{91}$$

Vector multiplication of Eq. (91) by $\boldsymbol{r}$ yields

$$\boldsymbol{r} \times \ddot{\boldsymbol{r}} = \frac{1}{\mu}\boldsymbol{r} \times \boldsymbol{f} = 0. \tag{92}$$

But the left-hand side of Eq. (92) can be developed [Eq. (4-41)] as

$$\boldsymbol{r} \times \ddot{\boldsymbol{r}} = \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{r} \times \dot{\boldsymbol{r}}) - (\dot{\boldsymbol{r}} \times \dot{\boldsymbol{r}}) = 0, \tag{93}$$

and thus,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{r} \times \dot{\boldsymbol{r}}) = 0 \tag{94}$$

and by integration,

$$\boldsymbol{r} \times \dot{\boldsymbol{r}} = \boldsymbol{C}. \tag{95}$$

The vector $\boldsymbol{C}$ is a constant which is perpendicular to the plane defined by $\boldsymbol{r}$ and the corresponding velocity $\dot{\boldsymbol{r}}$, and hence the momentum $\boldsymbol{p}$. This plane, which is called the collision plane, can be employed to describe the entire encounter between the two particles.

---

*Andres Jonas Ångström, Swedish physicist (1814–1871).

6.7.2 Conservation of energy

The kinetic energy of the hypothetical particle of mass μ in the collision plane, perpendicular to the z axis, can be expressed by

$$T = \tfrac{1}{2}\mu(\dot{x}^2 + \dot{y}^2), \tag{96}$$

or, in polar coordinates, with the substitutions $x = r\, cos\, \vartheta$ and $y = r\, sin\, \vartheta$, by

$$T = \tfrac{1}{2}\mu(\dot{r} + r^2\dot{\vartheta}^2). \tag{97}$$

The total energy is then the sum of Eq. (97) and the appropriate potential function for the particle interaction.

It is useful to define two parameters that, with the potential function, characterize the collision, namely,

(i) The impact parameter b, which is the distance of closest approach in the absence of the potential, and

(ii) The initial relative speed g of the colliding particles.

Before the advent of the collision ($r = \infty$) the potential is equal to zero and the kinetic energy $\tfrac{1}{2}\mu g^2$ is the total energy of the system. Furthermore, the angular momentum is given by μgb. Thus, the conservation of energy and angular momentum throughout the collision can be written as

$$\tfrac{1}{2}\mu g^2 = \tfrac{1}{2}\mu(\dot{r} + r^2\dot{\vartheta}^2) + V(r) \tag{98}$$

and

$$\mu bg = \mu r^2\dot{\vartheta}, \tag{99}$$

respectively, where the right-hand side of Eq. (99) is obtained by taking the partial derivative of Eq. (97) with respect to $\dot{\vartheta}$. Equations (98) and (99) can be combined by eliminating $\dot{\vartheta}$ to yield

$$\tfrac{1}{2}\mu g^2 = \tfrac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\mu g^2\left(\frac{b}{r}\right)^2 + V(r). \tag{100}$$

Integration of Eq. (100) allows r to be determined as a function of time; *i.e.* the trajectory of the collision can be specified if the potential function is known.

6.7.3 Interaction potential: LJ (6-12)

Many different empirical or semi-empirical functions have been suggested to represent the interaction between two spherical particles. The most successful is certainly that of Lennard-Jones.* Specifically, it is the Lennard-Jones (6-12) function, which has the form

$$V(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right], \tag{101}$$

where the parameter σ, an effective molecular diameter, is the value of r for which $V(r) = 0$. The minimum in the potential-energy curve occurs at $r = \sqrt[6]{\sigma}$, where its depth is equal to ε. The inverse sixth power of r represents the attractive forces which exist even between spherical particles. They are due to dispersion (London†) forces, as well as dipole–induced-dipole forces which are present when the particles are polar. The inverse twelfth function of the distance is an empirical representation of the repulsive forces, which increase rapidly at short distances. It is used for mathematical convenience. In general, the Lennard-Jones (6-12) potential function provides a useful and reliable representation of the interaction of atoms and nonpolar – or slightly polar – molecules.

A typical Lennard-Jones (6-12) function is plotted in Fig. 9. Often, the second term on the right-hand side of Eq. (100) is added to represent an effective potential function, *viz.*

$$V_{eff}(r) = V(r) + \tfrac{1}{2}\mu g^2\left(\frac{b}{r}\right)^2. \tag{102}$$

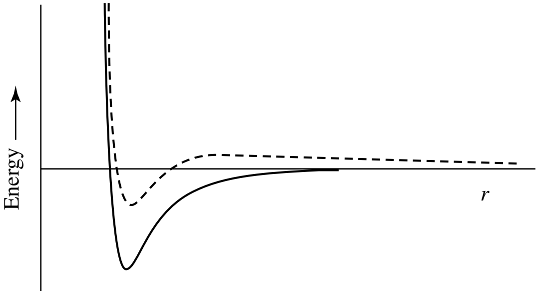


Fig. 9 Solid line: A typical Lennard-Jones (6-12) potential; dotted line: the effective potential for given values of the initial parameters g and b.

---

*John Edward Lennard-Jones, British theoretical physical chemist (1894–1954).

†Fritz London, German Physicist (1900–1954).

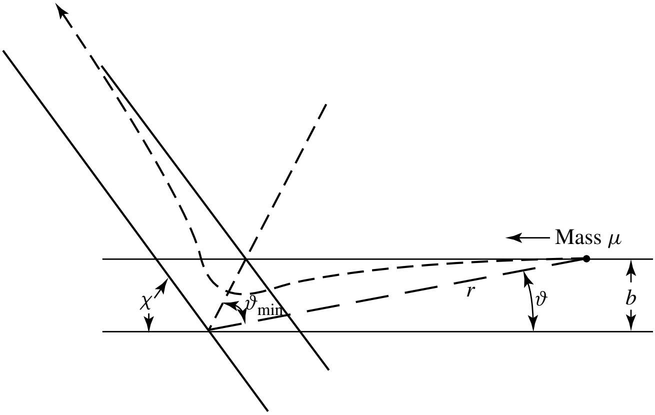


Fig. 10 Binary collision with respect to the center of mass of the system.

The dotted curve shown in Fig. 9 is an example, although it is but one in a family, as the effective potential depends on g and b, the parameters which define the initial conditions of a binary collision.

6.7.4 Angle of deflection

The result of a binary collision is specified in classical mechanics by the angle of deflection, χ. It is defined in Fig. 10, where the trajectory of the hypothetical particle of mass μ is illustrated. It can be seen that when r is minimal, that is at the distance of closest approach, the angle ϑ_{min} is related to the angle of deflection by

$$\chi = \pi - 2\vartheta_{min}. \tag{103}$$

From Eqs. (99) and (100), r and ϑ can be related by

$$\frac{\mathrm{d}r}{\mathrm{d}\vartheta} = \frac{\dot{r}}{\dot{\vartheta}} = -\left(\frac{r^2}{b}\right)\sqrt{1 - \frac{V(r)}{\frac{1}{2}\mu g^2} - \frac{b^2}{r^2}}; \tag{104}$$

the negative sign has been chosen so that r decreases with ϑ along the incoming trajectory. Thus, ϑ_{min} can be calculated from the expression

$$\vartheta_{min} = \int_0^{\vartheta_{min}} \mathrm{d}\vartheta = -b\int_\infty^{r_{min}} \frac{1}{\sqrt{1 - (b^2/r^2) - [V(r)/\frac{1}{2}\mu g^2]}}\frac{\mathrm{d}r}{r^2}, \tag{105}$$

which, with Eq. (103), allows the angle of deflection $\chi(b, g)$ to be determined.

For dilute, real gases, where ternary and higher collisions can be neglected, the angle of deflection can be employed to evaluate a number of physical properties. Of course appropriate distributions of the values of g and b must be introduced. The resulting expressions for the virial coefficients and the transport properties (viscosity, diffusion and thermal conductivity) are quite complicated. The interested reader is referred to advanced books on this subject.

6.7.5 Quantum mechanical description: The phase shift

In the classical picture of two-particle interaction outlined above, it was shown that a specific quantity – the angle of deflection – characterizes a given collision. However, on the atomic-molecular scale, quantum theory is more appropriate. According to the uncertainty principle of Heisenberg,* the simultaneous determination of the position and momentum of a particle cannot be made. Thus, it is not possible to determine exactly the angle of deflection in a collision. In the following development it is found that the phase shift of the radial wave function characterizes a binary, quantum-mechanical collision. This quantity, then, which is analogous to the classical angle of deflection, determines the final quantum-mechanical expressions describing the physical properties of low-pressure, real gases.

The radial equation for the quantum mechanical, two-particle system [Eq. (60)] can be applied to the present problem by employing the identity $\beta = \ell(\ell+1)$, as before, and making the substitution $R(r) = (1/r)S(r)$ used to obtain Eq. (72). The result is given by

$$\frac{\mathrm{d}^2 S(r)}{\mathrm{d}r^2} + \left\{ -\frac{\ell(\ell+1)}{r^2} + \frac{2\mu}{\hbar^2}[\varepsilon - V(r)] \right\} S(r) = 0. \tag{106}$$

In the case of an ideal gas, $V(r) = 0$ and Eq. (106) becomes

$$\frac{\mathrm{d}^2 S(r)}{\mathrm{d}r^2} + \left\{ -\frac{\ell(\ell+1)}{r^2} + \frac{\mu^2 g^2}{\hbar^2} \right\} S^0(r) = 0, \tag{107}$$

where the total energy has been equated to $\frac{1}{2}\mu g^2$. Equation (106) can be compared to the general form of Bessel's equation given in Section 5.5.5. It is in the present application

$$\frac{\mathrm{d}^2 S^0(r)}{\mathrm{d}r^2} + \left(\alpha^2 - \frac{4p^2 - 1}{4r^2} \right) S^0(r) = 0, \tag{108}$$

---

*Werner Heisenberg, German physicist (1901–1976).

Table 1 Characteristics of a binary collision.

Classical mechanics	Quantum mechanics
$\chi(b, g)$: Angle of deflection	$\eta_\ell(\alpha)$: Phase shift
b: Impact parameter	ℓ: Angular momentum quantum number
$\mu g b$: Angular momentum	$\hbar\sqrt{\ell(\ell+1)}$: Angular momentum
g: Initial relative speed	$\alpha = \mu g/\hbar = 2\pi/\lambda$: Wavenumber of the deBroglie wave
μg: Relative momentum	$\alpha\hbar$: Relative momentum

with the identifications $\alpha^2 = \mu^2 g^2/\hbar^2$ and $p = \pm(\ell + \frac{1}{2})$. The general solution to Bessel's equation is of the form

$$Z_p(\alpha r) = A_\ell(\alpha) J_{\ell+\frac{1}{2}}(ar) + B_\ell(\alpha) N_{-\ell-\frac{1}{2}}(\alpha r), \tag{109}$$

where $J_{\ell+\frac{1}{2}}(ar)$ and $N_{-\ell-\frac{1}{2}}(\alpha r)$ are the Bessel and Neumann functions, respectively. As the Neumann function becomes infinite in the limit as $r \to \infty$, the coefficient $B_\ell(\alpha)$ must be set equal to zero. Furthermore, the function $J_{\ell+\frac{1}{2}}(ar)$ becomes sinusoidal for large values of r, representing the deBroglie wave of the hypothetical particle.

In the more general problem in which $V(r) \neq 0$, the previous boundary condition is not applicable. Thus, $B_\ell(\alpha) \neq 0$ and the asymptotic solution for large values of r is given by [Eq. (5-148)]

$$S(r) = rR(r) = [A_\ell^2(\alpha) + B_\ell^2(\alpha)]^{1/2}\, sin\left[\alpha r - \tfrac{1}{2}\pi\ell + \eta_{\ell(\alpha)}\right]. \tag{110}$$

The argument of the sine in Eq. (110) now contains the phase shift,

$$\eta_\ell(\alpha) = tan^{-1}[(-1)^\ell B_\ell(\alpha)/A_\ell(\alpha)], \tag{111}$$

which represents the net result of the encounter. This quantity is analogous to the angle of deflection in the classical case.

The results of this section can be summarized by comparison with those of the previous one. Thus, the corresponding quantities in the classical and quantum-mechanical treatments of the collision problem are given in Table 1.

PROBLEMS

1. Derive Eqs. (18) and (20).

2. Derive Eq. (25).

3. Apply the initial conditions for the struck string (ii) to obtain the expression for $y(x, t)$.

4. Carry out the separation of variables to obtain Eq. (30).

5. Verify Eq. (36).

6. Show that the kinetic energy in the two-body problem in terms of momenta is given by Eq. (44).

7. Make the substitution $Y(\theta, \varphi) = \Theta(\theta)\Phi(\varphi)$ in Eq. (59) to obtain Eq. (61).

8. Verify the expression for a, as given by Eq. (77).

9. Calculate the error in the energy of the ground state of the hydrogen atom if the reduced mass of the two-particle system is replaced by the mass of the electron. Ans. 0.05%

10. Calculate the Rydberg constant from the values of the fundamental constants [see Eq. (84) and Appendix II]. Ans. $R_H = 109{,}737.5\ \text{cm}^{-1}$

11. Muonium, “atom number zero”, is composed of a positron and an electron. Calculate the Rydberg constant for this species.

Ans. $54{,}898.6\ \text{cm}^{-1}$

7 Operators and Matrices

The notion of an operator has already been developed and employed in Section 5.3. There, the operator $\hat{\mathcal{D}} \equiv \mathrm{d}/\mathrm{d}x$ was used in the solution of ordinary differential equations. In Chapter 5 the vector operator "del", represented by the symbol ∇, was introduced. It was shown that its algebraic form is dependent on the choice of curvilinear coordinates.

It is the objective of the present chapter to define matrices and their algebra – and finally to illustrate their direct relationship to certain operators. The operators in question are those which form the basis of the subject of quantum mechanics, as well as those employed in the application of group theory to the analysis of molecular vibrations and the structure of crystals.

7.1 THE ALGEBRA OF OPERATORS

The addition of operators follows the general rule of addition,

$$(\hat{\alpha} + \hat{\beta})f = \hat{\alpha}f + \hat{\beta}f; \tag{1}$$

that is, addition is distributive. Furthermore,

$$(\hat{\alpha} + \hat{\beta})f = (\hat{\beta} + \hat{\alpha})f. \tag{2}$$

Thus, addition is commutative.

The multiplication of two or more operators is accomplished by their successive application on a function. The order of the operations is by convention from right to left. Thus, the expression $\hat{\alpha}\hat{\beta}f$ implies that the operation $\hat{\beta}$ is carried out before the operation $\hat{\alpha}$. The multiplication of operators is associative, *viz.*

$$\hat{\alpha}\hat{\beta}f = \hat{\alpha}(\hat{\beta}f) = (\hat{\alpha}\hat{\beta})f. \tag{3}$$

However, it is not commutative, so that

$$\hat{\alpha}\hat{\beta}f \neq \hat{\beta}\hat{\alpha}f \tag{4}$$

in general.

As a simple example of the above principles, consider the operator $\hat{\mathcal{X}}$, defined as multiplication of the following function by the independent variable, say, x. Thus,

$$\hat{\mathcal{X}} f(x) = x f(x). \tag{5}$$

If a second operator is defined by the relation

$$\hat{\mathcal{D}} f(x) = \frac{\mathrm{d}f}{\mathrm{d}x} = f'(x), \tag{6}$$

as in Section 4.3, the products of these operators can be evaluated as

$$\hat{\mathcal{D}}\hat{\mathcal{X}} f(x) = \hat{\mathcal{D}}(\hat{\mathcal{X}} f) = \hat{\mathcal{D}}(x\ f) = x f' + f \tag{7}$$

and

$$\hat{\mathcal{X}}\hat{\mathcal{D}} f(x) = \hat{\mathcal{X}}(\hat{\mathcal{D}} f) = \hat{\mathcal{X}} f' = x f'. \tag{8}$$

The difference yields the relation

$$(\hat{\mathcal{D}}\hat{\mathcal{X}} - \hat{\mathcal{X}}\hat{\mathcal{D}}) f(x) = f(x). \tag{9}$$

The corresponding operator relation is

$$\hat{\mathcal{D}}\hat{\mathcal{X}} - \hat{\mathcal{X}}\hat{\mathcal{D}} = 1. \tag{10}$$

However, such a relation is meaningless unless it is understood that the operators are followed by a function upon which they operate.

The commutation relations involving operators are expressed by the so-called commutator, a quantity which is defined by

$$\hat{\alpha}\hat{\beta} - \hat{\beta}\hat{\alpha} = \left[\hat{\alpha}, \hat{\beta}\right]. \tag{11}$$

Thus, for the example presented above, the commutator is given by

$$\left[\hat{\mathcal{D}}, \hat{\mathcal{X}}\right] = 1 \tag{12}$$

and the operators $\hat{\mathcal{D}}$ and $\hat{\mathcal{X}}$ do not commute. This result is of fundamental importance in quantum mechanics, as will be demonstrated at the end of this chapter.

The operators that are involved in quantum mechanics are linear. An example of a linear operator is given by

$$\hat{\alpha}[c_1 f_1(x) + c_2 f_2(x)] = c_1 \hat{\alpha} f_1(x) + c_2 \hat{\alpha} f_2(x). \tag{13}$$

On the other hand, if an operator $\hat{\beta}$ is defined by

$$\hat{\beta} f(x) = [f(x)]^2, \tag{14}$$

it should be apparent that

$$\hat{\beta} f_1(x) + \hat{\beta} f_2(x) \neq \hat{\beta}[f_1(x) + f_2(x)]. \tag{15}$$

Such an operator is nonlinear, and it will not appear in quantum-mechanical applications.

So far, nothing has been said concerning the nature of the functions, such as $f(x)$ in the above examples, upon which the operators operate. In practical problems the functions are said to be "well behaved"! This expression means that the functions are:

(i) continuous
(ii) single-valued
(iii) finite.*

These restrictions are in general the origin of the boundary conditions imposed on the solutions of the Schrödinger equation, as illustrated in Chapter 5.

7.2 HERMITIAN OPERATORS AND THEIR EIGENVALUES

It is important to note that all operators of interest in quantum mechanics are Hermitian (or self-adjoint). This property is defined by the relation

$$\int \eta^{\star}(\hat{\alpha}\xi)\,\mathrm{d}\tau = \int \xi(\hat{\alpha}^{\star}\eta^{\star})\,\mathrm{d}\tau, \tag{16}$$

where $\hat{\alpha}$ is an operator and the functions η and ξ are well-behaved, as defined above. The importance of this property will become more apparent in later sections of this chapter.

As an example, consider the quantum mechanical operator for the linear momentum in one dimension,

$$p_x \longrightarrow \frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}x} = \hat{p}_x,$$

which was employed in Section 6.3.2. It, and the coordinate x are mutually conjugate, as illustrated earlier. The Hermitian property follows from

---

*This condition is too severe, as it is the integral $\int f^{\star} f\,\mathrm{d}\tau$ that must remain finite.

the relations

$$\int \eta^\star(x)\left(\frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}x}\right)\xi(x)\,\mathrm{d}x = -i\hbar\int \eta^\star\frac{\mathrm{d}\xi}{\mathrm{d}x}\,\mathrm{d}x \tag{17}$$

$$= -i\hbar\eta^\star\xi\Big|_{-\infty}^{+\infty} + i\hbar\int \xi\frac{\mathrm{d}}{\mathrm{d}x}\eta^\star\,\mathrm{d}x$$

$$= i\hbar\int \xi\frac{\mathrm{d}}{\mathrm{d}x}\eta^\star\,\mathrm{d}x. \tag{18}$$

The first term in Eq. (18) is equal to zero at each limit for the assumed well-behaved functions. Thus, Eqs. (17) and (18) lead to

$$\int \eta^\star(x)\hat{p}\xi(x)\,\mathrm{d}x = \int \xi(x)\hat{p}^\star\eta^\star(x)\,\mathrm{d}x, \tag{19}$$

in agreement with Eq. (16).

The characteristic-value problem – more often referred to as the eigenvalue problem – is of extreme importance in many areas of physics. Not only is it the very basis of quantum mechanics, but it is employed in many other applications. Given a Hermitian operator $\hat{\alpha}$, if their exists a function (or functions) ς such that

$$\hat{\alpha}\varsigma = a\varsigma, \tag{20}$$

the values of a are known as the eigenvalues of $\hat{\alpha}$. The functions ς are the corresponding eigenfunctions.

It is important to note that the eigenvalues of a Hermitian operator are real. If Eq. (20) is multiplied by $\varsigma^\star$ and the integration is carried out over all space, the result is

$$\int \varsigma^\star\hat{\alpha}\varsigma\,\mathrm{d}\tau = a\int \varsigma^\star\varsigma\,\mathrm{d}\tau. \tag{21}$$

The complex conjugate of Eq. (20) can be written as

$$\hat{\alpha}^\star\varsigma^\star = a^\star\varsigma^\star. \tag{22}$$

Multiplication of Eq. (22) by ς and integration over all space yields the relation

$$\int \varsigma\hat{\alpha}^\star\varsigma^\star\,\mathrm{d}\tau = a^\star\int \varsigma\varsigma^\star\,\mathrm{d}\tau. \tag{23}$$

As $\hat{\alpha}$ is Hermitian, the left-hand side of Eq. (21) is equal to the left-hand side of Eq. (23). Therefore,

$$a\int \varsigma^\star\varsigma\,\mathrm{d}\tau = a^\star\int \varsigma\varsigma^\star\,\mathrm{d}\tau \tag{24}$$

and $a = a^\star$. Thus, the eigenvalues of Hermitian operators are real. It can be shown that the inverse is true. Since the eigenvalues correspond to physically observable quantities, they are real and their operators are Hermitian.

As an example, consider Eq. (6-61), which can be written as

$$\frac{\mathrm{d}^2}{\mathrm{d}\varphi^2}\Phi = -m^2\Phi. \tag{25}$$

In this form it can be compared to Eq. (20), where the result of the operation $\mathrm{d}^2/\mathrm{d}\varphi^2$ is to multiply the function Φ by the eigenvalues $-m^2$. Clearly, the eigenfunctions are of the form $\Phi = e^{im\varphi}$. As shown in Section 6.4.2, the quantization of m, *viz.* $m = 0, \pm 1, \pm 2, \ldots$ is the result of the conditions imposed on the solutions. In this case it is the requirement that they be single-valued. It will be shown in Section 9.2.4 that the operator $\hat{L}_z = (\hbar/i)(\mathrm{d}/\mathrm{d}\varphi)$ corresponds to the z component of the angular momentum of the system. That is, for the hydrogen atom it is the vertical component of the angular momentum of the electron. Equation (25) is then equivalent to

$$\hat{L}_z^2\Phi = -\hbar^2\frac{\mathrm{d}^2}{\mathrm{d}\varphi^2}\Phi = m^2\hbar^2\Phi. \tag{26}$$

7.3 MATRICES

A matrix is an array of numbers. For most practical purposes it is rectangular. Thus, a matrix is an array such as

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \tag{27}$$

where the elements a_{ij} are numbers or functions, which may be real or complex. The subscripts i and j of the element a_{ij} identify the row and the column, respectively, of the matrix in which it is located. The matrix given in Eq. (27) consists of m rows and n columns. If the matrix is square, $m = n$, and the matrix is said to be of order n. In a square matrix of order n the elements $a_{11}, a_{22}, \ldots, a_{nn}$ constitute the main diagonal of $\boldsymbol{A}$. The sum of the diagonal elements of a matrix is called the trace (German: *Spur*). In group theory it is known as the character, *i.e.* the quantity that characterizes a matrix representation (see Chapter 8). If all of the nondiagonal elements in a matrix are equal to zero, the matrix is diagonal.

Two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are equal if they are identical. That is, $a_{ij} = b_{ij}$ for every pair of subscripts. The addition of two matrices can be defined by the sum $\boldsymbol{A} + \boldsymbol{B}$, which is the matrix of elements $[(a_{ij} + b_{ij})]_{mn}$. That is, the sum of two matrices of the same order is found by adding their corresponding elements. Note that the two matrices must be congruent, that is, have the same number of rows and columns. The product of a number c and a matrix $\boldsymbol{A}$ is defined as the matrix $\boldsymbol{B}$ whose elements are the elements of $\boldsymbol{A}$ multiplied by c. Namely, $\boldsymbol{B} = c\boldsymbol{A}$ if $b_{ij} = ca_{ij}$.

The multiplication of matrices requires a bit more reflection. The product $\boldsymbol{C}$ of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is usually defined by $\boldsymbol{C} = \boldsymbol{AB}$ if

$$c_{ij} = \sum_k a_{ik} b_{kj}. \tag{28}$$

The sum in Eq. (28) is over the number of columns of $\boldsymbol{A}$, which must of course be equal to the number of rows of $\boldsymbol{B}$. The result is the matrix $\boldsymbol{C}$, whose number of rows is equal to the number of rows of $\boldsymbol{A}$ and whose number of columns is equal to the number of columns of $\boldsymbol{B}$.

It is important to note that the product of two square matrices, given by $\boldsymbol{AB}$ is not necessarily equal to $\boldsymbol{BA}$. In other words, matrix multiplication is not commutative. However, the trace of the product does not depend on the order of multiplication. From Eq. (28) it is apparent that

$$Tr(\boldsymbol{AB}) = \sum_i (\boldsymbol{AB})_{ii} = \sum_i \sum_k a_{ik} b_{ki} = \sum_k \sum_i b_{ki} a_{ik} = Tr(\boldsymbol{BA}). \tag{29}$$

While the matrix multiplication defined by Eq. (28) is the more usual one in matrix algebra, there is another way of taking the product of two matrices. It is known as the direct product and is written here as $\boldsymbol{A} \otimes \boldsymbol{B}$. If $\boldsymbol{A}$ is a square matrix of order n and $\boldsymbol{B}$ is a square matrix of order m, then $\boldsymbol{A} \otimes \boldsymbol{B}$ is a square matrix of order nm. Its elements consist of all possible pairs of elements, one each from $\boldsymbol{A}$ and $\boldsymbol{B}$, *viz.*

$$[\boldsymbol{A} \otimes \boldsymbol{B}]_{ik,jl} = a_{ij} b_{kl}. \tag{30}$$

The arrangement of the elements in the direct-product matrix follows certain conventions. They are illustrated in the following chapter, where the direct product of matrices is employed in the theory of groups.

The unit matrix $\boldsymbol{E}$ (German: *Einheit*) is one which is diagonal with all of the diagonal elements equal to one. It plays the role of unity in matrix algebra. Clearly, the unit matrix multiplied by a constant yields a diagonal matrix with all of the diagonal elements equal to the value of the constant. If the constant is equal to zero, the matrix is the null matrix $\boldsymbol{0}$, with all elements equal to zero.

In many applications in physics and chemistry there appear systems of linear equations of the general form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned} \tag{31}$$

where the number m of equations is not necessarily equal to the number n of unknowns. Following the definition of matrix multiplication, as given in Eq. (28), this system of equations can be written as the matrix equation

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}. \tag{32}$$

Then, with $\boldsymbol{A} = [a_{ij}], \boldsymbol{X} = [x_j]$ and $\boldsymbol{B} = [b_i]$ Eq. (32) becomes simply

$$\boldsymbol{AX} = \boldsymbol{B}. \tag{33}$$

The bold-face characters employed in Eq. (33) imply that each symbol represents a matrix. The problem of the resolution of simultaneous linear equations will be discussed in Section 7.8, as certain properties of matrices must first be explained.

The matrices such as $\boldsymbol{X}$ and $\boldsymbol{B}$ in Eq. (33), which are composed of a single column, are usually referred to as vectors. In fact, the vectors introduced in Chapter 4 can be written as column matrices in which the elements are the corresponding components. Of course the vector $\boldsymbol{X} = [x_j]$ in Eq. (33) is of dimension n, while those in Chapter 4 were in three-dimensional space. It is apparent that the matrix notation introduced here is a more general method of representing vector algebra in multidimensional spaces. This idea is developed further in Section 7.7.

The transpose of a matrix $\boldsymbol{A}$ is a matrix $\tilde{\boldsymbol{A}}$, which is obtained from $\boldsymbol{A}$ by interchanging rows and columns, *viz.* $[\tilde{\boldsymbol{A}}]_{ij} = a_{ji}$. Clearly, the transpose of a column matrix, a vector, is a row matrix. The complex conjugate of $\boldsymbol{A}$ is the matrix $\boldsymbol{A}^{\bigstar}$ whose elements are the complex conjugates of the corresponding elements of $\boldsymbol{A}$. The conjugate transpose is the matrix $\boldsymbol{A}^{\dagger}$, which is the complex conjugate of the transpose of $\boldsymbol{A}$; that is, $\boldsymbol{A}^{\dagger} = \tilde{\boldsymbol{A}}^{\bigstar}$.

The inverse, $\boldsymbol{A}^{-1}$, of a matrix $\boldsymbol{A}$ is defined by the relation $\boldsymbol{AA}^{-1} = \boldsymbol{E}$. If $\boldsymbol{A}$ is a square matrix, its inverse may exist – although not necessarily so. This question is addressed later in this section. Rectangular, nonsquare, matrices may

Table 1 Special matrices.

Description	Condition	Elements
Real	$\boldsymbol{A}^\star = \boldsymbol{A}$	$a_{ij}^\star = a_{ij}$
Symmetric	$\tilde{\boldsymbol{A}} = \boldsymbol{A}$	$a_{ji} = a_{ij}$
Hermitian	$\boldsymbol{A}^\dagger = \boldsymbol{A}$ or $\boldsymbol{A}^\star = \tilde{\boldsymbol{A}}$	$a_{ji}^\star = a_{ij}$
Orthogonal	$\boldsymbol{A}^{-1} = \tilde{\boldsymbol{A}}$ or $\boldsymbol{A}\tilde{\boldsymbol{A}} = \boldsymbol{E}$	$\left[\boldsymbol{A}^{-1}\right]_{ji} = a_{ij}$
Unitary	$\boldsymbol{A}^{-1} = \boldsymbol{A}^\dagger$ or $\boldsymbol{A}\boldsymbol{A}^\dagger = \boldsymbol{E}$	$\left[\boldsymbol{A}^{-1}\right]_{ji}^\star = a_{ij}$

also possess inverses, although this question is somewhat more complicated (see problem 4).

Several special matrices are defined in Table 1. The Hermitian matrix is of particular importance in quantum-mechanical applications, as outlined in Section 7.13.

It is often necessary to take the transpose of a product of matrices. Thus, if $\boldsymbol{AB} = \boldsymbol{C}$, $\boldsymbol{c}_{ij} = \sum_k a_{ik}b_{kj}$, where in the general case all three matrices are rectangular [see Eq. (28)]. If both $\boldsymbol{A}$ and $\boldsymbol{B}$ are transposed, their product is taken in the order $\tilde{\boldsymbol{B}}\tilde{\boldsymbol{A}}$, as the number of columns of $\tilde{\boldsymbol{B}}$ must be equal to the number of rows of $\tilde{\boldsymbol{A}}$. The result is the transpose of $\boldsymbol{C}$, namely, $\tilde{\boldsymbol{C}} = \tilde{\boldsymbol{B}}\tilde{\boldsymbol{A}}$. This principle holds for any number of factors; thus, when a matrix product is transposed, the sequence of the matrices forming the product must be reversed, *e.g.*

$$\boldsymbol{F} = \boldsymbol{ABC}\ldots\boldsymbol{X}, \quad \tilde{\boldsymbol{F}} = \tilde{\boldsymbol{X}}\ldots\tilde{\boldsymbol{C}}\tilde{\boldsymbol{B}}\tilde{\boldsymbol{A}}. \tag{34}$$

A similar relation applies to the inverse of the product of matrices. For example, define the product of three matrices by $\boldsymbol{ABC} = \boldsymbol{W}$. If the inverse of $\boldsymbol{W}$ exists, it is given by $\boldsymbol{W}^{-1} = (\boldsymbol{ABC})^{-1}$. Now consider the product

$$(\boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1})\boldsymbol{ABC} = (\boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1})\boldsymbol{W}, \tag{35}$$

where it is assumed that the inverse of each matrix, $\boldsymbol{A}$ $\boldsymbol{B}$ and $\boldsymbol{C}$ exists. As the associative law holds and $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{E}$, *etc.*, the left-hand side of Eq. (35) is equal to the identity; then,

$$\boldsymbol{E} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}\boldsymbol{W} \tag{36}$$

and

$$\boldsymbol{W}^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1} = (\boldsymbol{ABC})^{-1}. \tag{37}$$

This result can be easily generalized to include the inverse of the product of any number of factors.

7.4 THE DETERMINANT

For most students, their first encounter with matrices is in the study of determinants. However, a determinant is a very special case in which a given square matrix has a specific numerical value. If the matrix $\boldsymbol{A}$ is of order two, its determinant can be written in the form

$$|\boldsymbol{A}| = Det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \equiv \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}. \tag{38}$$

For a matrix of order three, its determinant can be developed in the form

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \tag{39}$$

Clearly, each determinant of order two in Eq. (39) can be evaluated following Eq. (38). These determinants of order two are called the minors of the determinant of order three. Note that the minor of a_{11} is obtained by elimination of the row and column in which it appears. Similarly, that of the element a_{12} is obtained by eliminating its row and column. Furthermore, the second term in Eq. (39) is negative because the sum of the subscripts of a_{12} is odd. The sign is positive if the sum of the subscripts is even. A minor with its appropriate sign is referred to as a cofactor. The principles outlined in this paragraph are general and can thus be applied to determinants of higher order.

Although the development of determinants of any order can be made, as illustrated in Eq. (39), in the special case of matrices of third order there is another, often useful, method. It is shown in Fig. 1. The solid arrows, starting with elements a_{11}, a_{12} and a_{13} pass through elements which form the products $a_{11}a_{22}a_{33}$, $a_{21}a_{32}a_{13}$ and $a_{31}a_{23}a_{12}$, respectively. Similarly, following the dotted arrows, the products $a_{31}a_{22}a_{13}$, $a_{21}a_{12}a_{33}$ and $a_{11}a_{23}a_{32}$ are obtained.

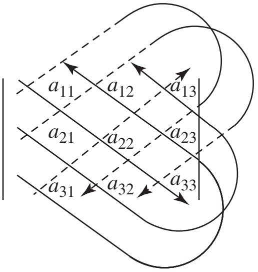


Fig. 1 The development of a third-order determinant.

Addition of these six products, with negative signs for those obtained from the dotted arrows, yields the result

$$\begin{aligned}|\boldsymbol{A}| = {} & a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{23}a_{12} \\ & - (a_{31}a_{22}a_{13} + a_{21}a_{12}a_{33} + a_{11}a_{23}a_{32}),\end{aligned} \tag{40}$$

which can be easily verified by comparison with Eq. (39). It must be emphasized that the method illustrated in Fig. 1 is only applicable to determinants of order three.

The matrix represented in this chapter by $\overline{\boldsymbol{A}}$ is usually called the adjoint matrix. It is obtained by constructing the matrix which is composed of all of the cofactors of the elements a_{ij} in $|\boldsymbol{A}|$ and then taking its transpose. With the basic definition of matrix multiplication [Eq. (29)] and some patience, the reader can verify the relation

$$\boldsymbol{A}\overline{\boldsymbol{A}} = \overline{\boldsymbol{A}}\boldsymbol{A} = |\boldsymbol{A}|\boldsymbol{E} \tag{41}$$

(see problem 8). If the determinant $|\boldsymbol{A}|$ is equal to zero, $\boldsymbol{A}$ is said to be singular and $\boldsymbol{A}\overline{\boldsymbol{A}} = \overline{\boldsymbol{A}}\boldsymbol{A} = \boldsymbol{0}$. If $\boldsymbol{A}$ is nonsingular, Eq. (41) can be divided by $|\boldsymbol{A}|$ to yield a matrix $\boldsymbol{A}^{-1}$ which is the inverse of $\boldsymbol{A}$, or,

$$\boldsymbol{A}^{-1} = \frac{\overline{\boldsymbol{A}}}{|\boldsymbol{A}|} \tag{42}$$

and $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{E}$. Thus, the inverse of a square matrix exists only if it is nonsingular.

7.5 PROPERTIES OF DETERMINANTS

Some general properties of determinants can be summarized as follows.

(i) The value of a determinant is unchanged if its rows and columns are interchanged, *viz.*

$$|\boldsymbol{A}| = |\tilde{\boldsymbol{A}}|. \tag{43}$$

(ii) The sign of a determinant changes when two rows (or columns) are interchanged, *e.g.*

$$\begin{vmatrix} a_{12} & a_{11} & a_{13} \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \tag{44}$$

as is easily shown by expansion of the two determinants.

(iii) If the elements of a given row (or column) are multiplied by the same quantity, say, c, it can be removed as a common factor, *viz.*

$$\begin{vmatrix} ca_{11} & ca_{12} & ca_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = c \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}. \tag{45}$$

(iv) If two rows (or columns) of a determinant are identical, the value of the determinant is zero; thus,

$$\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0. \tag{46}$$

(v) The product of two determinants is equal to the determinant of the matrix product of the two, *e.g.*

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = \begin{vmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{vmatrix}. \tag{47}$$

The proof of (v) constitutes problem 7.

7.6 JACOBIANS

Partial derivatives, as introduced in Section 2.12 are of particular importance in thermodynamics. The various state functions, whose differentials are exact (see Section 3.5), are related *via* approximately 10^{10} expressions involving 720 first partial derivatives! Although some of these relations are not of practical interest, many are. It is therefore useful to develop a systematic method of deriving them. The method of Jacobians is certainly the most widely applied to the solution of this problem. It will be only briefly described here. For a more advanced treatment of the subject and its application to thermodynamics, the reader is referred to specialized texts.

Consider two functions $x(u, v)$ and $y(u, v)$, where u and v are independent variables. In this case the Jacobian can be defined by the determinant

$$J(x, y|u, v) \equiv \begin{vmatrix} \left(\dfrac{\partial x}{\partial u}\right)_v & \left(\dfrac{\partial x}{\partial v}\right)_u \\ \left(\dfrac{\partial y}{\partial u}\right)_v & \left(\dfrac{\partial y}{\partial v}\right)_u \end{vmatrix}, \tag{48}$$

whose expansion yields

$$J(x, y|u, v) = \left(\frac{\partial x}{\partial u}\right)_v \left(\frac{\partial y}{\partial v}\right)_u - \left(\frac{\partial x}{\partial v}\right)_u \left(\frac{\partial y}{\partial u}\right)_v. \tag{49}$$

A number of significant properties of the Jacobian can be easily derived:

$$\begin{aligned}
&\text{(i)} & J(x,v|u,v) &= (\partial x/\partial u)_v \\
&\text{(ii)} & J(x,y|x,y) &= 1 \\
&\text{(iii)} & J(x,x|u,v) &= 0 \\
&\text{(iv)} & J(x,c|u,v) &= 0, \text{ where } c \text{ is a constant.} \\
&\text{(v)} & J(x,y|u,v) &= -J(x,y|v,u) \\
&\text{(vi)} & J(x,y|u,v) &= -J(y,x|u,v) \\
&\text{(vii)} & J(x,y|u,v)J(u,v|s,t) &= J(x,y|s,t) \\
&\text{(viii)} & J(x,y|u,v) &= 1/[J(u,v|x,y)].
\end{aligned} \tag{50}$$

Relation (i) is evident, as

$$\left(\frac{\partial v}{\partial u}\right)_v = 0 \text{ and } \left(\frac{\partial v}{\partial v}\right)_u = 1.$$

It allows any partial derivative to be expressed as a Jacobian. Equations (iii) to (vi) follow directly from the properties of the determinant described in the previous section, while (ii) and (viii) are the result of the general expressions for partial derivatives (Section 2.12). Finally, relation (vii) requires a bit more thought. With the use of the definition of the Jacobian and property (v) of determinants [Eq. (50)], the 1-1 element of the resulting Jacobian is

$$\left(\frac{\partial x}{\partial u}\right)_v \left(\frac{\partial u}{\partial s}\right)_t + \left(\frac{\partial x}{\partial v}\right)_u \left(\frac{\partial v}{\partial s}\right)_t = \left(\frac{\partial x}{\partial s}\right)_t, \tag{51}$$

where the chain rule has been applied. The other elements can be found in a similar manner.

In Section 2.13 it was shown that for bulk systems the various thermodynamic functions are related by the system of equations

$$\begin{aligned}
T &= \left(\frac{\partial E}{\partial S}\right)_V = \left(\frac{\partial H}{\partial S}\right)_P, \ P = -\left(\frac{\partial E}{\partial V}\right)_S = -\left(\frac{\partial F}{\partial V}\right)_T \\
V &= \left(\frac{\partial H}{\partial P}\right)_S = \left(\frac{\partial G}{\partial P}\right)_T, \ S = -\left(\frac{\partial F}{\partial T}\right)_V = -\left(\frac{\partial G}{\partial T}\right)_P.
\end{aligned} \tag{52}$$

In the Jacobian notation these relations become

$$\begin{aligned}
T &= J(E,V|S,V) = J(H,P|S,P), \\
P &= J(E,S|S,V) = J(F,T|T,V), \\
V &= J(H,S|P,S) = J(G,T|P,T),
\end{aligned}$$

and

$$S = J(F, V|V, T) = J(G, P|P, T). \tag{53}$$

In this notation Maxwell's relations take the form

$$J(T, S|V, S) = J(P, V|V, S),$$

$$J(T, S|P, S) = J(P, V|P, S),$$

$$J(T, S|T, V) = J(P, V|T, V),$$

and

$$J(T, S|T, P) = J(P, V|T, P). \tag{54}$$

As an example of the use of Jacobians to obtain thermodynamic relationships, consider the quantity

$$\left(\frac{\partial V}{\partial T}\right)_S = J(V, S|T, S) = \frac{J(V, S|T, V)}{J(T, S|T, V)}, \tag{55}$$

where T and V are taken as independent variables and rule (vii) has been used to obtain the second equality. With use of rule (vi) and Maxwell's relations,

$$\left(\frac{\partial V}{\partial T}\right)_S = -\frac{J(S, V|T, V)}{J(T, S|T, V)} = -\frac{J(S, V|T, V)}{J(P, V|T, V)}. \tag{56}$$

Finally, as

$$\left(\frac{\partial S}{\partial T}\right)_V = \frac{C_V}{T}$$

under adiabatic and reversible conditions,

$$\left(\frac{\partial V}{\partial T}\right)_S = -\frac{C_V}{T\left(\dfrac{\partial P}{\partial T}\right)_V}. \tag{57}$$

The partial derivative is now expressed in terms of the heat capacity and the equation of state, which are experimental quantities.

7.7 VECTORS AND MATRICES

As indicated above, if the components of a vector $\boldsymbol{X}$ in n-dimensional space are real, the vector can be written as a column matrix with n rows. Similarly, a second vector $\boldsymbol{Y}$ in the same space can be written as a column matrix of the same order. The scalar product of these two vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ written in matrix notation, becomes

$$\tilde{\boldsymbol{X}}\boldsymbol{Y} = \sum_{i=1}^{n} X_i Y_i, \tag{58}$$

where in Eq. (5-12) $n = 3$. The result given in Eq. (58) should be evident if it is understood that the transpose of a column matrix is a row matrix.

When $n > 3$ in Eq. (58), the space cannot be visualized. However, the analogy with three-dimensional space is clear. Thus an n-dimensional coordinate system consists of n mutually perpendicular axes. A point requires n coordinates for its location and, which is equivalent, a vector is described by its n components.

If $\boldsymbol{A}$ is a square matrix and $\boldsymbol{X}$ is a column matrix, the product $\boldsymbol{AX}$ is also a column. Therefore, the product $\tilde{\boldsymbol{X}}\boldsymbol{AX}$ is a number. This matrix expression, which is known as a quadratic form, arises often in both classical and quantum mechanics (Section 7.13). In the particular case in which $\boldsymbol{A}$ is Hermitian, the product $\boldsymbol{X}^{\dagger}AX$ is called a Hermitian form, where the elements of $\boldsymbol{X}$ may now be complex.

The vector product $\boldsymbol{X} \times \boldsymbol{Y}$ is somewhat more complicated in matrix notation. In the three-dimensional case, an antisymmetric (or skew symmetric) matrix can be constructed from the elements of the vector $\boldsymbol{X}$ in the form

$$\boldsymbol{X} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}. \tag{59}$$

The vector product is then obtained by ordinary matrix multiplication,

$$\begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{bmatrix}. \tag{60}$$

The column matrix on the left-hand side of Eq. (60) is a vector, whose elements are the coefficients of $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$, respectively, in the notation of Chapter 5. Thus,

$$\boldsymbol{X} \times \boldsymbol{Y} = \begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \end{bmatrix} \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \end{bmatrix}, \tag{61}$$

as in Eq. (4-24).

7.8 LINEAR EQUATIONS

In general, Eq. (32) represents a system of inhomogeneous linear equations. It is assumed that $\boldsymbol{A}$ and $\boldsymbol{B}$ are known and the elements of the vector $\boldsymbol{X}$ are the unknowns. For simplicity, the following arguments will be limited to the case in which $\boldsymbol{A}$ is square, that is, $n = m$. If all elements of the vector $\boldsymbol{B}$ are equal to zero, the equations are homogeneous and Eq. (33) becomes $\boldsymbol{AX} = \boldsymbol{0}$.

The solution of a system of linear equations depends on certain conditions, *viz.*

(i) If $\boldsymbol{B} \neq \boldsymbol{0}$ and $|\boldsymbol{A}| \neq 0$, then $\boldsymbol{A}^{-1}$ exists and the unique solutions are given by

$$\boldsymbol{X} = \boldsymbol{A}^{-1}\boldsymbol{B} = \frac{\overline{\boldsymbol{A}}\boldsymbol{B}}{|\boldsymbol{A}|}. \tag{62}$$

(ii) If $\boldsymbol{B} = \boldsymbol{0}$ and $|\boldsymbol{A}| \neq 0$, the only solution is the trivial one, with $\boldsymbol{X} = \boldsymbol{0}$; that is, all of the unknowns are equal to zero.

(iii) However, if $\boldsymbol{B} = \boldsymbol{0}$, a nontrivial solution to these homogeneous equations exists if $|\boldsymbol{A}| = 0$. This condition is usually referred to as Cramer's rule.* It should be noted, however, that because the equations are homogeneous, only ratios of the unknowns can be evaluated. Thus, an additional relation among the unknowns must be invoked in order to obtain unique solutions. This problem is of great importance in many applications, in particular in the classical theory of molecular vibrations and in quantum mechanics. It will be developed in more detail in Chapter 9.

7.9 PARTITIONING OF MATRICES

It is often useful to partition matrices, either square or rectangular, into submatrices, as indicated by the following examples:

$$\boldsymbol{A} = \left[\begin{array}{ccc:cc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ \hdashline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{array}\right] = \left[\begin{array}{c:c} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \hdashline \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{array}\right], \tag{63}$$

where the bold-faced letters with subscripts identify the corresponding submatrices. Thus, with the use of one or more dashed lines a matrix can be

---

*Gabriel Cramer, Swiss mathematician (1704–1752).

partitioned into submatrices whose positions in the original matrix are specified by the subscripts.

The addition of two or more partitioned matrices is straightforward, providing of course that they are partitioned in the same way. Then, for example,

$$\boldsymbol{A}+\boldsymbol{B}=\left[\begin{array}{c:c}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \hdashline \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{array}\right]+\left[\begin{array}{c:c}\boldsymbol{B}_{11} & \boldsymbol{B}_{12}\\ \hdashline \boldsymbol{B}_{21} & \boldsymbol{B}_{22}\end{array}\right]$$

$$=\left[\begin{array}{c:c}\boldsymbol{A}_{11}+\boldsymbol{B}_{11} & \boldsymbol{A}_{12}+\boldsymbol{B}_{12}\\ \hdashline \boldsymbol{A}_{21}+\boldsymbol{B}_{21} & \boldsymbol{A}_{22}+\boldsymbol{B}_{22}\end{array}\right]. \tag{64}$$

The product of partitioned matrices can be obtained in a similar manner if the columns of $\boldsymbol{A}$ match the rows of $\boldsymbol{B}$, *e.g.*

$$\boldsymbol{A}\boldsymbol{B}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \boldsymbol{A}_{13}\end{bmatrix}\begin{bmatrix}\boldsymbol{B}_{11}\\ \boldsymbol{B}_{12}\\ \boldsymbol{B}_{22}\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_{11}\boldsymbol{B}_{11} & \boldsymbol{A}_{12}\boldsymbol{B}_{12} & \boldsymbol{A}_{13}\boldsymbol{B}_{13}\end{bmatrix}. \tag{65}$$

This result should be obvious from the definition of matrix multiplication [Eq. (28)].

7.10 MATRIX FORMULATION OF THE EIGENVALUE PROBLEM

The eigenvalue problem was introduced in Section 7.3, where its importance in quantum mechanics was stressed. It arises also in many classical applications involving coupled oscillators. The matrix treatment of the vibrations of polyatomic molecules provides the quantitative basis for the interpretation of their infrared and Raman spectra.* This problem will be addressed more specifically in Chapter 9.

The eigenvalue problem can be described in matrix language as follows. Given a matrix $\boldsymbol{H}$, determine the scalar quantities λ and the nonzero vectors $\boldsymbol{L}_i$ which satisfy simultaneously the equation

$$\boldsymbol{H}\boldsymbol{L}_i=\lambda\boldsymbol{L}_i. \tag{66}$$

In all physical applications, although both $\boldsymbol{H}$ and $\boldsymbol{L}_i$ may contain complex elements, the eigenvalues λ, are real (see Section 7.2). Equation (66) can be

---

*Sir Chandrasekhara Venkata Raman, Indian physicist (1886–1970).

written in the form

$$(\boldsymbol{H} - \lambda\boldsymbol{E})\boldsymbol{L}_i = \mathbf{0}. \tag{67}$$

If the unit matrix $\boldsymbol{E}$ is of order n, Eq. (67) represents a system of n homogeneous, linear equations in n unknowns. They are usually referred to as the secular equations. According to Cramer's rule [see (iii) of Section 7.8], nontrivial solutions exist only if the determinant of the coefficients vanishes. Thus, for the solutions of physical interest,

$$|\boldsymbol{H} - \lambda\boldsymbol{E}| = 0; \tag{68}$$

that is, the secular determinant vanishes. It is perhaps useful to write out Eq. (68) as,

$$\begin{vmatrix} h_{11} - \lambda & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} - \lambda & \cdots & h_{2n} \\ \vdots & \vdots & & \\ h_{n1} & h_{n2} & \cdots & h_{nn} - \lambda \end{vmatrix} = 0. \tag{69}$$

This relation is equivalent to an algebraic equation of degree n in the unknown λ and therefore has n roots, some of which may be repeated (degenerate). These roots are the characteristic values or eigenvalues of the matrix $\boldsymbol{H}$.

When the determinant of Eq. (69) is expanded, the result is the polynomial equation

$$(-\lambda)^n + c_1(-\lambda)^{n-1} + c_2(-\lambda)^{n-2} + \ldots + c_n = 0. \tag{70}$$

The coefficients c_i are given by

$$c_1 = \sum_i h_{ii} = Tr\boldsymbol{H} = \sum_i \lambda_i,$$

$$c_2 = \sum_{j,i<j} (h_{ii}h_{jj} - h_{ij}h_{ji}) = \sum_{j,i<j} \lambda_i\lambda_j,$$

$$\vdots$$

and

$$c_n = |\boldsymbol{H}| = \prod_{i=1}^{n} \lambda_i. \tag{71}$$

The equalities in Eqs. (71) are the result of a well-known theorem of algebra. As $\boldsymbol{H}$ is nonsingular, the expression for c_n shows that no eigenvalue can be equal to zero.

Now if λ_1 is one of the eigenvalues of $\boldsymbol{H}$, Eq. (66) is satisfied and the equation

$$(\boldsymbol{H} - \lambda_1 \boldsymbol{E})\boldsymbol{L}_1 = \boldsymbol{0} \tag{72}$$

has nontrivial solutions for the vector $\boldsymbol{L}_1$. If n eigenvectors $\boldsymbol{L}_i$ have been found, they may be assembled into a square matrix $\boldsymbol{L} = [\boldsymbol{L}_1 \vdots \boldsymbol{L}_2 \vdots \boldsymbol{L}_3 \vdots \cdots \vdots \boldsymbol{L}_n]$. Then, Eq. (66) becomes

$$\boldsymbol{HL} = \boldsymbol{L\Lambda} \tag{73}$$

or

$$\boldsymbol{L}^{-1}\boldsymbol{HL} = \boldsymbol{\Lambda}, \tag{74}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix whose elements are the eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n$ of $\boldsymbol{H}$. From Eq. (74) it is apparent that the eigenvalues are obtained by diagonalization of $\boldsymbol{H}$ by a similarity transformation, *e.g.* premultiplication by $\boldsymbol{L}^{-1}$, followed by postmultiplication by $\boldsymbol{L}$. A transformation of this type, the similarity transformation, will take on a considerable importance in the applications of group theory presented in the following chapter.

7.11 COUPLED OSCILLATORS

A simple eigenvalue problem can be demonstrated by the example of two coupled oscillators. The system is illustrated in Fig. 2. It should be compared with the classical harmonic oscillator that was treated in Section 5.2.2. Here also, the system will be assumed to be harmonic, namely, that both springs obey Hooke's law. The potential energy can then be written in the form

$$V = \tfrac{1}{2}\kappa_1 x_1^2 + \tfrac{1}{2}\kappa_2 (x_2 - x_1)^2, \tag{75}$$

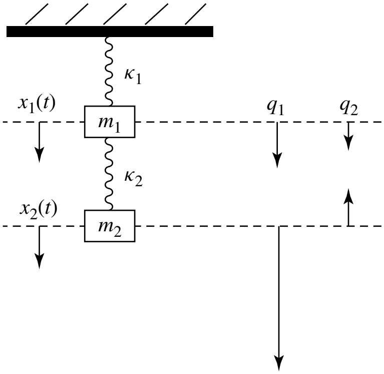


Fig. 2 System of two coupled oscillators.

where κ_1 and κ_2 are the force constants. The corresponding expression for the kinetic energy is given simply by

$$T = \tfrac{1}{2} m_1 \dot{x}_1^2 + \tfrac{1}{2} m_2 \dot{x}_2^2. \tag{76}$$

Then, applying Newton's law in the form

$$\frac{\mathrm{d}}{\mathrm{d}t} p_i = f_i, \tag{77}$$

where $p_i = m_i \dot{x}_i$ and f_i is the force acting on mass m_i. In terms of the potential and kinetic energies given by Eqs. (75) and (76), respectively,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{x}_i}\right) = -\frac{\partial V}{\partial x_i}, \tag{78}$$

and the equations of motion become

$$m_1 \ddot{x}_1 + \kappa_1 x_1 - \kappa_2 (x_2 - x_1) = 0 \tag{79}$$

and

$$m_2 \ddot{x}_2 + \kappa_2 (x_2 - x_1) = 0. \tag{80}$$

The introduction of periodic solutions of the form

$$x_i = a_i\, cos[\sqrt{\lambda} t + \eta], \tag{81}$$

leads to the pair of simultaneous equations

$$\frac{a_1(\kappa_1 + \kappa_2)}{m_1} - \frac{a_2 \kappa_2}{m_1} = a_1 \lambda \tag{82}$$

and

$$-\frac{a_1 \kappa_2}{m_2} + \frac{a_2 \kappa_2}{m_2} = a_2 \lambda \tag{83}$$

for the unknown amplitudes a_1 and a_2. It should be noted that $\sqrt{\lambda}$ can be identified with $\omega = 2\pi\nu^0$, the angular frequency. In matrix form Eqs. (82) and (83) can be written as

$$\begin{bmatrix} \dfrac{\kappa_1 + \kappa_2}{m_1} & -\dfrac{\kappa_2}{m_1} \\ -\dfrac{\kappa_2}{m_2} & \dfrac{\kappa_2}{m_2} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{84}$$

or

$$\begin{bmatrix} \dfrac{\kappa_1+\kappa_2}{m_1}-\lambda & -\dfrac{\kappa_2}{m_1} \\ -\dfrac{\kappa_2}{m_2} & \dfrac{\kappa_2}{m_2}-\lambda \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \boldsymbol{0}, \tag{85}$$

which is a system of linear, homogeneous equations. According to Cramer's rule, for a nontrivial solution to be obtained, the determinant of the coefficients must vanish; then,

$$\begin{vmatrix} \dfrac{\kappa_1+\kappa_2}{m_1}-\lambda & -\dfrac{\kappa_2}{m_1} \\ -\dfrac{\kappa_2}{m_2} & \dfrac{\kappa_2}{m_2}-\lambda \end{vmatrix} = 0, \tag{86}$$

leading to a quadratic equation for the two eigenvalues. This result corresponds to the general form given by Eq. (68).

For simplicity, the following development of this example will be limited to the special case in which $m_1 = m_2 = m$ and $\kappa_1 = \kappa_2 = \kappa$. Then, the expansion of the secular determinant of Eq. (86) yields

$$\lambda^2 - \frac{3\kappa}{m}\lambda + \frac{\kappa^2}{m^2} = 0, \tag{87}$$

whose roots are

$$\lambda = \frac{\kappa}{2m}(3 \pm \sqrt{5}). \tag{88}$$

By choosing arbitrarily one of the roots as the eigenvalue λ_1 and substituting it in the secular equations, a relation between the amplitudes can be obtained, namely,

$$a_{21} = \tfrac{1}{2}(1+\sqrt{5})a_{11}, \tag{89}$$

where the second subscript has been added to specify that this relation between the amplitudes was obtained with the use of the first, arbitrarily chosen eigenvalue.

As pointed out above, the homogeneous secular equations can only yield expressions for the ratios of the amplitudes. The additional condition necessary to resolve this ratio is the normalization, which in this example is

$$a_{21}^2 + a_{11}^2 = 1. \tag{90}$$

In general, the normalization of the amplitudes is very useful, as will be indicated below – and developed in more detail in the treatment of molecular vibrations (see Chapter 9). The normalized amplitudes obtained by combination of Eqs. (89) and (90) are $\ell_{11} = \sqrt{2}(5+\sqrt{5})^{-1/2}$ and $\ell_{21} = (1+\sqrt{5})[2(5+\sqrt{5})]^{-1/2}$. The repetition of the procedure employed in the last paragraph with the use of the second eigenvalue results in the normalized amplitudes $\ell_{12} = \sqrt{2}(5 - \sqrt{5})^{-1/2}$ and $\ell_{22} = (1 - \sqrt{5})[2(5 - \sqrt{5})]^{-1/2}$. Thus, the matrix $\boldsymbol{L}$ takes the form

$$\boldsymbol{L} = \begin{bmatrix} \sqrt{2}(5 + \sqrt{5})^{-1/2} & \sqrt{2}(5 - \sqrt{5})^{1/2} \\ (1 + \sqrt{5})[2(5 + \sqrt{5})]^{-1/2} & (1 - \sqrt{5})[2(5 - \sqrt{5})]^{-1/2} \end{bmatrix}. \tag{91}$$

It should be noted that this matrix is orthogonal, *viz.* $\tilde{\boldsymbol{L}} = \boldsymbol{L}^{-1}$ and thus, $\tilde{\boldsymbol{L}}\boldsymbol{L} = \boldsymbol{E}$.

With a bit of patience it can be verified that Eq. (74) becomes in this case

$$\boldsymbol{L}^{-1}\boldsymbol{H}\boldsymbol{L} = \tilde{\boldsymbol{L}}\boldsymbol{H}\boldsymbol{L} = \boldsymbol{\Lambda}, \tag{92}$$

where

$$\boldsymbol{H} = \begin{bmatrix} 2\kappa/m & -\kappa/m \\ -\kappa/m & \kappa/m \end{bmatrix} \tag{93}$$

and

$$\boldsymbol{\Lambda} = \begin{bmatrix} \frac{\kappa}{2m}(3 - \sqrt{5}) & 0 \\ 0 & \frac{\kappa}{2m}(3 + \sqrt{5}) \end{bmatrix} \tag{94}$$

(see problem 11).

A quantitative physical picture of the vibrations of the system illustrated in Fig. 2 can be obtained with the use of normal coordinates. However, it is first customary to introduce the so-called mass-weighted coordinates by $s_i = \sqrt{m_i}x_i$, where of course $m_1 = m_2 = m$ in this example. The normal coordinates $\boldsymbol{Q}$ can then be defined by the relation

$$\boldsymbol{S} = \begin{bmatrix} \sqrt{m}x_1 \\ \sqrt{m}x_2 \end{bmatrix} = \boldsymbol{L}\begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \boldsymbol{L}\boldsymbol{Q}. \tag{95}$$

They have the property of allowing both the potential and the kinetic energies to be written as sums of square terms; that is, all cross terms vanish. This result can be demonstrated by substituting for x_1 and x_2 in Eq. (95) into Eqs. (75) and (76). The results are

$$V = \tfrac{1}{2}(\lambda_1 q_1^2 + \lambda_2 q_2^2) \tag{96}$$

and

$$T = \tfrac{1}{2}(\dot{q}_1^2 + \dot{q}_2^2) \tag{97}$$

(see problem 12).

As $\boldsymbol{L}$ is orthogonal, the normal coordinates can be found from Eq. (95) as

$$\boldsymbol{Q} = \tilde{\boldsymbol{L}}\boldsymbol{S} = \sqrt{m}\tilde{\boldsymbol{L}}\boldsymbol{X}. \tag{98}$$

The relative displacements of the masses in the two normal modes of this coupled oscillator are shown to the right in Fig. 2. This method of representing the form of the normal modes is particularly useful in the analysis of molecular vibrations (see Chapter 9).

7.12 GEOMETRIC OPERATIONS

The trivial geometric operation is known as the identity. If it is applied to an arbitrary vector in, say, three-dimensional space, the result is to leave the vector ξ unchanged. In Cartesian coordinates this operation can be expressed in matrix form as

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tag{99}$$

The unit matrix is a representation of the trivial operation of "doing nothing". Equation (99) can be written more compactly as $\hat{\boldsymbol{E}}\xi = \xi$. The circumflex over $\boldsymbol{E}$ indicates that it is an operator, although in practice this mark is often omitted.

Other operations on a vector include reflections. Consider, for example, the operation of reflection in the x,y plane. The result of this operation is to change the sign of the z component of the vector. Thus, a reflection in the x, y plane can be represented by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ -z \end{bmatrix}. \tag{100}$$

This operation can be considered to be an inversion of the single coordinate z, as shown in the following chapter. The symbol $\hat{\boldsymbol{\sigma}}(xy)$, which is often used for this operation, is that of Schönflies.* Clearly, the other two reflections in the Cartesian planes correspond to the matrix relations

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ -y \\ z \end{bmatrix} \tag{101}$$

---

*Arthur Moritz Schönflies, German mathematician (1853–1928). The Schönflies symbols are employed in spectroscopic applications, while in crystallography the international, or Hermann-Maugin notation is used.

and

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x \\ y \\ z \end{bmatrix}, \tag{102}$$

with the symbols $\hat{\boldsymbol{\sigma}}(zx)$ and $\hat{\boldsymbol{\sigma}}(yz)$, respectively.

The operation which changes the signs of all three components of $\boldsymbol{\xi}$ is called the inversion. Its representation is obviously a negative unit matrix. Then,

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x \\ -y \\ -z \end{bmatrix}. \tag{103}$$

The Schönflies notation for this operation is $\hat{\imath}$, although the circumflex is usually omitted.

To find the representation of a simple or "proper" rotation it is convenient to choose the rotation axis along a Cartesian direction. For a rotation of the coordinates by an arbitrary angle about, say, the z axis, that component of any vector will be unchanged. However, the x and y components will be converted into linear combinations, as can be seen by reference to Fig. 3. Here, for simplicity, a unit vector $\boldsymbol{1}$ has been chosen. When the coordinates undergo a counter-clockwise rotation about the z axis by an angle φ the components of the unit vector become $x' = cos(\alpha - \varphi) = cos\,\alpha\, cos\,\varphi + sin\,\alpha\, sin\,\varphi$ and $y' = sin(\alpha - \varphi) = -\,cos\,\alpha\, sin\,\varphi + sin\,\alpha\, cos\,\varphi$. In matrix notation the result is

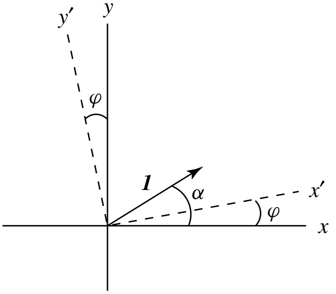


Fig. 3 The effect of a counter-clockwise coordinate rotation about the z axis.

expressed in the form*

$$\begin{bmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}. \tag{104}$$

The Schönflies symbol for a rotation by an angle φ is $C_{2\pi/\varphi}$. If, for example, $\varphi = \pi/2$, the symbol becomes C_4 (see the following chapter).

If a proper rotation is combined with a reflection with respect to the axis of rotation, it is called an improper rotation The matrix representation of such an operation is found simply by replacing 1 by -1 in Eq. (104). The Schönflies symbol for an improper rotation by φ is $S_{2\pi/\varphi}$. Hence, matrix the representation of an improper counter-clockwise rotation by φ is of the form†

$$S_{2\pi/\varphi} = \begin{bmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{105}$$

It should be noted that the trace of a matrix that represents a given geometric operation is equal to $2\cos\varphi \pm 1$, the choice of signs is appropriate to proper or improper operations. Furthermore, it should be noted that the arbitrary direction of rotation has no effect on the value of the trace, as a rotation in the inverse sense corresponds only to a change in sign of the element $\sin\varphi$. These operations and their matrix representations will be employed in the following chapter, where the theory of groups is applied to the analysis of molecular symmetry.

7.13 THE MATRIX METHOD IN QUANTUM MECHANICS

In classical mechanics the kinetic energy of a particle of mass m is written in the form

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2), \tag{106}$$

where $\dot{x}$, $\dot{y}$ and $\dot{z}$ are the Cartesian components of its velocity. The corresponding momentum components are given by

$$p_x = \frac{\partial T}{\partial \dot{x}} = m\dot{x}, \tag{107}$$

---

*The linear transformation expressed by Eq. (104) has the same form if the vector is rotated in the ***clockwise*** direction by the angle φ while the coordinate axes remain fixed.

†The Schönflies symbols are usually not written in bold-face type, although they are represented by matrices such as given in Eqs. (104) and (105).

and similarly in the y and z directions. Substitution for the velocity components in Eq. (106) yields the kinetic energy expression

$$T = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) = \frac{\boldsymbol{p}^2}{2m}. \tag{108}$$

In quantum mechanics the vector $\boldsymbol{p}$ is replaced by the operator $(\hbar/i)\nabla$. Thus, the operator for the kinetic energy becomes

$$\hat{T} = \frac{\hat{\boldsymbol{p}}^2}{2m} = -\frac{\hbar^2}{2m}\nabla^2. \tag{109}$$

The operator for the total energy of the system is then obtained by adding the potential energy, V, which is a function of the appropriate coordinates. The result, which is known as the Hamiltonian, is then

$$\hat{H} = \hat{T} + V = -\frac{\hbar^2}{2m}\nabla^2 + V. \tag{110}$$

This principle was employed earlier (see Sections 5.4 and 6.3.2).

The operators of interest in quantum mechanics such as the Hamiltonian, as well as the linear and angular momenta, are Hermitian (see Section 7.2). Their eigenvalues are then real, as they correspond to quantities that are measurable.

Another fundamental property concerns the orthogonality of eigenfunctions. Given that

$$\hat{\alpha} f_1 = a_1 f_1 \text{ and } \hat{\alpha} f_2 = a_2 f_2, \tag{111}$$

the relation

$$\int f_1^\star \hat{\alpha} f_2 \, \mathrm{d}\tau = a_2 \int f_1^\star f_2 \, \mathrm{d}\tau \tag{112}$$

is easily established. And, as a_1 is real and $\hat{\alpha}$ is Hermitian,

$$\int f_1^\star \hat{\alpha} f_2 \, \mathrm{d}\tau = \int f_2 \hat{\alpha}^\star f_1^\star \, \mathrm{d}\tau = a_1^\star \int f_1^\star f_2 \, \mathrm{d}\tau = a_1 \int f_1^\star f_2 \, \mathrm{d}\tau. \tag{113}$$

Therefore, the difference, Eqs. (112) minus (113) yields

$$(a_2 - a_1) \int f_1^\star f_2 \, \mathrm{d}\tau = 0, \tag{114}$$

and either $a_2 - a_1 = 0$ or $\int f_1^\star f_2 \, \mathrm{d}\tau = 0$, or both. In the former case the system is degenerate, as two eigenvalues are the same; or, if $\int f_1^\star f_2 \, \mathrm{d}\tau = 0$, the functions f_1 and f_2 are orthogonal. It can be concluded, then, that the eigenfunctions in a nondegenerate system are orthogonal. However, it can be shown that in the case of a degenerate system, orthogonal eigenfunctions can be constructed as linear combinations.

A basic theorem of quantum mechanics, which will be presented here without proof, is: If $\hat{\alpha}$ and $\hat{\beta}$ commute, namely $[\hat{\alpha}, \hat{\beta}] = 0$, there exists an ensemble of functions that are eigenfunctions of both $\hat{\alpha}$ and $\hat{\beta}$ – and inversely.

In many applications of quantum mechanics in physics and chemistry, interest is primarily in the description of the stationary, or time-independent, states of a system. Thus, it is sufficient to determine the energies and wavefunctions with the use of the Schrödinger equation in the form

$$\hat{H}\psi_n = \varepsilon_n \psi_n. \tag{115}$$

Equation (115) is a special case of the more general Schrödinger equation which includes the time as an independent variable (see Section 12.3). The function ψ_n, is the eigenfunction for the state n with corresponding energy ε_n. Since the eigenfunctions are orthonormal (or can be made so), Eq. (115) can be multiplied by $\psi_m^\star$ and integrated over all space. The result is given by

$$\int \psi_m^\star \hat{H} \psi_n \, \mathrm{d}\tau = \varepsilon_n \delta_{m,n}, \tag{116}$$

where $\mathrm{d}\tau$ is the element of volume and $\delta_{m,n}$ is the Kronecker delta [see Eq. (5-105)]. The integrals on the left-hand side of Eq. (116) can be arranged in the form of a square matrix $\boldsymbol{H}$, whose order is equal to the number of independent eigenfunctions in the set. As the functions ψ are in this case the eigenfunctions for the problem, the matrix $\boldsymbol{H}$ is diagonal and equal to the right-hand side of Eq. (116), whose elements are the eigenvalues – that is, the values of the energies of the different stationary states. The diagonal matrix composed of elements $\varepsilon_n \; \delta_{m,n}$ is often represented by the symbol $\boldsymbol{\Lambda}$, as in Eq. (73).

If another set of orthonormal functions, say $\chi = \chi_1, \chi_2, \chi_3, \ldots$ is used to calculate $\boldsymbol{H}$, the result is a matrix composed of elements $\int \chi_i^\star \hat{H} \chi_j \, \mathrm{d}\tau$. It will not in general be in diagonal form. However, the set of functions χ can serve as a basis for the expression of the unknown wavefunctions. In general, a given function such as ψ_n can be expanded in a complete set of orthonormal functions. A Fourier series such as Eq. (6-15) is an example. If the complete set of orthonormal (and often complex) functions χ is available, the wavefunction appearing in Eq. (115) can be written as

$$\psi_n = \sum_j c_j \chi_j, \tag{117}$$

where the summation extends over all members of the set.* As each function χ_j is assumed to be known, and orthonormal in a given region, a knowledge of all of the coefficients c_j is sufficient to evaluate the correct eigenfunctions. If these coefficients are now written as a vector, namely,

$$\zeta_n = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix} \tag{118}$$

for each state n. These eigenvectors are collected as described in Section 7.10 to form a square matrix ζ. Then, the eigenvalue problem is that defined by Eqs. (73) and (74), namely,

$$\boldsymbol{H}\zeta = \zeta\boldsymbol{\Lambda}, \tag{119}$$

or

$$\zeta^{-1}\boldsymbol{H}\zeta = \boldsymbol{\Lambda}. \tag{120}$$

As the matrix ζ is unitary, $\zeta^{-1} = \zeta^{\dagger} = \tilde{\zeta}^{\star}$.

7.14 THE HARMONIC OSCILLATOR

The harmonic oscillator has already been presented in several forms. It is defined by the potential function $V = \frac{1}{2}\kappa\xi^2$ in one dimension. The corresponding quantum-mechanical problem, which leads to the wavefunctions presented in Table 5-1, yielded the expression for the energy: $\varepsilon = h\nu^0\left(v + \frac{1}{2}\right) = \hbar\omega(v + \frac{1}{2})$, with $v = 0, 1, 2, \ldots$ [Eq. (5-92)].

The energy of the one-dimensional, classical harmonic oscillator can be written in the form

$$H = T + V = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2, \tag{121}$$

where, following conventions, q is the independent variable and p is its conjugate momentum. The quantum-mechanical operators do not commute, as $[\hat{p}, \hat{q}] = -i\hbar$ (problem 13). This result is easily generalized, namely,

$$\hat{p}_j\hat{q}_k - \hat{q}_k\hat{p}_j = -i\hbar\delta_{j,k} \tag{122}$$

in more than one dimension.

---

*If the functions χ are continuous in a given region, the summation in Eq. (117) implies integration over that region.

In the matrix formulation of quantum mechanics Eq. (121) is transformed. If $\boldsymbol{P}$ and $\boldsymbol{Q}$ are matrices, the Hamiltonian becomes by analogy

$$\boldsymbol{H}(\boldsymbol{P},\boldsymbol{Q}) = \frac{1}{2m}(\boldsymbol{P}^2 + m^2\omega^2\boldsymbol{Q}^2). \tag{123}$$

If it is assumed that the matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ have been chosen so that $[\boldsymbol{P},\boldsymbol{Q}] = -i\hbar\boldsymbol{E}$, a given element of $\boldsymbol{H}$ is of the form

$$H_{j,k} = \frac{1}{2m}(\boldsymbol{P}^2 + m^2\omega^2\boldsymbol{Q}^2)_{j,k} = \varepsilon_j\delta_{j,k}. \tag{124}$$

Thus, the Hamiltonian has been diagonalized and the diagonal elements correspond to the energies of the system.

To obtain explicit expressions for the energies, define two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ by

$$\boldsymbol{A} = \boldsymbol{P} - im\omega\boldsymbol{Q}, \tag{125}$$

and

$$\boldsymbol{B} = \boldsymbol{P} + im\omega\boldsymbol{Q}. \tag{126}$$

Their products are given by

$$\boldsymbol{AB} = 2m\boldsymbol{H} + m\omega\hbar\boldsymbol{E} \tag{127}$$

and

$$\boldsymbol{BA} = 2m\boldsymbol{H} - m\omega\hbar\boldsymbol{E}. \tag{128}$$

The matrix product $\boldsymbol{ABA}$ is then equal to

$$\boldsymbol{A}(2m\boldsymbol{H} + m\omega\hbar\boldsymbol{E}) = (2m\boldsymbol{H} - m\omega\hbar\boldsymbol{E})\boldsymbol{A} \tag{129}$$

or,

$$\sum_i A_{k,i}\left(\varepsilon_i\delta_{i,j} - \frac{1}{2}\omega\hbar\delta_{i,j}\right) = \sum_i(\varepsilon_k\delta_{k,i} + \omega\hbar\delta_{k,i})A_{i,j}. \tag{130}$$

Thus,

$$A_{k,j}(\varepsilon_j - \varepsilon_k - \omega\hbar) = 0 \tag{131}$$

and $A_{k,j}$ vanishes unless

$$\varepsilon_j - \varepsilon_k = \hbar\omega. \tag{132}$$

The same result can be obtained from the product $\boldsymbol{BAB}$.

From Eq. (128)

$$(BA)_{jj} = \sum_k B_{j,k} A_{k,j} = 2m(\varepsilon_j - \tfrac{1}{2}\omega\hbar). \tag{133}$$

However, each term in the sum over k vanishes unless Eq. (132) is satisfied. Then, for a particular value of ε_j, either $\varepsilon_j - \varepsilon_k = \hbar\omega$ or there is no eigenvalue below ε_j, by an amount $\hbar\omega$. In the latter case the sum in Eq. (133) vanishes and $\varepsilon_j = \frac{1}{2}\hbar\omega$, which is the lowest eigenvalue. It can be concluded from this argument that the eigenvalues form a sequence, $\frac{1}{2}\hbar\omega$, $\frac{3}{2}\hbar\omega$, $\frac{5}{2}\hbar\omega$ *etc.*, in agreement with the values given by Eq. (5-92).

It should be emphasized that the development presented in this section yielded the same result for the energies of the harmonic oscillator as obtained by application of the Schrödinger equation. However, the notion of a wavefunction is absent. The matrix approach, as invented by Heisenberg, provides a direct method of obtaining the values of measurable quantities in a given system – in this case the energies of a one-dimensional harmonic oscillator. It must be admitted that the matrix formulation is more abstract than Schrödinger's method and its success often depends on judicious guesses.

PROBLEMS

1. Given the matrix

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 5 \\ 0 & 1 & 2 \end{pmatrix}$$

(a) What is the value of the trace of $\boldsymbol{A}$? Ans. 7

(b) Calculate the value of the determinant: $|\boldsymbol{A}|$. Ans. 9

(c) What are the values of the determinants $\begin{vmatrix} 1 & 2 & 0 \\ 3 & 5 & 4 \\ 0 & 2 & 1 \end{vmatrix}$ and $\begin{vmatrix} 1 & 2 & 0 \\ 3 & 5 & 0 \\ 0 & 2 & 0 \end{vmatrix}$? Ans. $-9, 0$

(d) Find the inverse of $\boldsymbol{A}$. Ans. $\begin{pmatrix} \frac{1}{3} & \frac{2}{9} & -\frac{8}{9} \\ -\frac{2}{3} & \frac{2}{9} & \frac{1}{9} \\ \frac{1}{3} & -\frac{1}{9} & \frac{4}{9} \end{pmatrix}$

2. Given the vector $\boldsymbol{B} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$,

(a) Calculate the product $\boldsymbol{AB} = \boldsymbol{C}$, where $\boldsymbol{A}$ is given in problem 1.

Ans. $\begin{pmatrix} 7 \\ 26 \\ 8 \end{pmatrix}$

(b) Evaluate the product $\tilde{\boldsymbol{B}}\boldsymbol{A}$ (where $\tilde{\boldsymbol{B}}$ is the transpose of $\boldsymbol{B}$). Ans. $\begin{pmatrix} 7 \\ 11 \\ 18 \end{pmatrix}$

3. Given the system of simultaneous equations

$$\begin{aligned} 3v - w + 2x + 6y &= 0 \\ 2v + 3w + 5x + 4y &= 0 \\ 4v - 2w + 2x + 7y &= 0 \\ -3v + 2w - x - 3y &= 0 \end{aligned}$$

Does a nontrivial solution exist? Explain. Ans. Yes, Section 7.8

4. Find an inverse of the matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

Ans. $\begin{pmatrix} 1 & 1 \\ -4 & -1 \\ \frac{8}{3} & \frac{1}{3} \end{pmatrix}$, *etc.*

5. Calculate the eigenvalues of the matrix $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Ans. $\dfrac{5 \pm \sqrt{33}}{2}$

6. Given the matrix: $\boldsymbol{H} = \begin{pmatrix} 1 & \dfrac{\sqrt{7}}{2} \\ \dfrac{\sqrt{7}}{2} & 1 \end{pmatrix}$:

(a) Calculate the eigenvalues λ_1 and λ_2. Ans. $\dfrac{9}{2}, \dfrac{1}{2}$

(b) Find the normalized eigenvectors $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$. Ans. $\begin{pmatrix} \dfrac{1}{\sqrt{8}} \\ \sqrt{\dfrac{7}{8}} \end{pmatrix}, \begin{pmatrix} -\sqrt{\dfrac{7}{8}} \\ \dfrac{1}{\sqrt{8}} \end{pmatrix}$

(c) To verify the results of (a) and (b), show that $\boldsymbol{HL} = \boldsymbol{L\Lambda}$, where

$$\boldsymbol{L} = [\boldsymbol{L}_1|\boldsymbol{L}_2] \text{ and } \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

(d) Find the inverse $\boldsymbol{L}^{-1}$. Ans. $\begin{pmatrix} \dfrac{1}{\sqrt{8}} & \sqrt{\dfrac{7}{8}} \\ -\sqrt{\dfrac{7}{8}} & \dfrac{1}{\sqrt{8}} \end{pmatrix}$

(e) Show that $\boldsymbol{H}$ is diagonalized by the similarity transformation, *viz.*

$$\boldsymbol{L}^{-1}\boldsymbol{H}\boldsymbol{L} = \boldsymbol{\Lambda}.$$

7. Prove property (v) of determinants (Section 7.5).

8. Verify Eq. (41).

9. Derive the properties of Jacobians, as given by Eqs. (50).

10. Verify Eq. (57).

11. Verify Eqs. (93) and (94).

12. Verify Eqs. (96) and (97).

13. Show that $[\hat{p}, \hat{q}] = -i\hbar$.

8 Group Theory

A group is a set of abstract elements (members) that has specific mathematical properties. In general it is not necessary to specify the nature of the members of the group or the way in which they are related. However, in the applications of group theory of interest to physicists and chemists, the key word is symmetry.

It is interesting that one of the earliest notions of what might be called structural chemistry is that of Plato.* His conception of matter was in terms of geometrical forms. The "elements" were then volumes, limited by surfaces in the form of triangles. Thus, Plato envisioned three regular polyhedra, the tetrahedron, the octahedron and the icosahedron, formed from equilateral triangular surfaces. The fourth polyhedron, constructed from twelve right, isosceles triangles was the cube. These polyhedra were thought to represent the "chemical" elements: fire, air, water and earth, respectively.

Students of modern chemistry know that the atoms in molecules and crystals are usually arranged in an ordered manner. It is therefore possible to describe the geometry of a molecule as, say, planar, tetrahedral, *etc.* Perhaps the most obvious application of group theory is its use in specifying such geometries. As will be shown in this chapter, it is equally capable of describing the symmetry of mathematical functions. And, it is hoped that the equivalence of these two applications will be made clear. As an historical note, the basis of this branch of mathematics was established by Galois.†

8.1 DEFINITION OF A GROUP

For a set of abstract elements to form a group in the mathematical sense, the following conditions must be satisfied.

(i) The combination of any two members of the set is an element of the set.
(ii) There is one element of the set which permutes with all of the others and leaves them unchanged. It is known as the identity.

---

*Plato, Greek philosopher (427–348 B.C.).

†Evariste Galois, French mathematician (1811–1832), who died at the age of twenty in a duel.

(iii) The associative law of combination is valid.
(iv) Each element in the set has an inverse that is also a member of the set.

If the elements are designated by $X, Y, Z \ldots$, they are said to belong to the group $\mathcal{G}$. Thus, condition (i) can be stated: If $X \in \mathcal{G}$ and $Y \in \mathcal{G}$, then $XY \in \mathcal{G}$. Note that the combination of X and Y as written here suggests multiplication of the elements X and Y. However, the word "multiplication" is taken to mean the law of combination, which must be specified.

If the identity is represented by the symbol E, condition (ii) can be written: $E \in \mathcal{G}$ such that $EX = XE = X, \forall X$. Clearly, E plays the role of unity in group theory and it is not by chance that the symbol is the same as that used for the unit matrix (see Section 7.3).

Condition (iii) is written as $X(YZ) = (XY)Z$. It was stated for operators in the previous chapter [Eq. (7-3)].

The existence of the inverse or reciprocal can be stated as follows: If $X \in \mathcal{G}$, $X^{-1} \in \mathcal{G}$ such that $X^{-1}X = XX^{-1} = E$.

For a given application of group theory it is furthermore necessary to define both the set of elements $X, Y, Z \ldots$ and the law of combination ("multiplication" or other). The following examples should clarify this question. It is to be noted that the order of the group is defined as the number of its members.

8.2 EXAMPLES

An infinite group can be formed which consists of all whole numbers (positive, negative and zero) and for which the combination law is ordinary addition. The identity is $E = 0$, as for any whole number $n + 0 = n$. In this example, then, the inverse $n^{-1} = -n$, because $n + (-n) = 0$.

As a second example, consider the group formed by the elements 1, $i, -1, -i$, where $i^2 = -1$. These elements are developed by the operations i^n, where n is an integer. If the law of combination is ordinary multiplication, the multiplication table for this cyclic group can be developed (Table 1). This

Table 1 Multiplication table for the cyclic group of order four.

	$E \equiv i^0 = 1$	i	$i^2 = -1$	$-i$
$i^0 = 1$	1	i	-1	$-i$
i	i	-1	$-i$	1
$i^2 = -1$	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

Table 2 Multiplication table for the four-group composed of the matrices of Eqs. (1).

	E	A	B	C
E	E	A	B	C
A	A	E	C	B
B	B	C	E	A
C	C	B	A	E

table is seen to be symmetric with respect to the principal diagonal. Thus, multiplication is in this case commutative and the group is said to be Abelian.*

Now consider the group composed of the following diagonal matrices:

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1}$$

If the combination law is matrix multiplication as defined in Eq. (7-28), the multiplication table shown as Table 2 can be constructed. This group is known as the "four-group" (German: *Vierergruppe*), the dihedral group of order four. It is apparent from its multiplication table that it is also Abelian. It should be noted that all of the square matrices that form this group [Eqs. (1)] are nonsingular, that is, their determinants are not equal to zero. This condition is necessary for the existence of their inverses [See (iv) above and Section 7.4]. All other groups of fourth order are isomorphic† with one or the other of the two groups defined above.

As a somewhat more complicated example, consider the group of order six composed of the following nonsingular matrices:

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}, \quad D = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \text{ and } \quad F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2}$$

It is not too difficult to develop the multiplication table shown as Table 3 (problem 1). It will be noticed immediately that the table is not symmetric with respect to the principal diagonal. Therefore, the group is not Abelian and multiplication is not commutative.

---

*After the Norwegian mathematician Niels Abel (1802–1829), who died of tuberculosis at the age of 26.

†Two groups are isomorphic [Greek, $\iota\sigma o$- (same or equal) $+\mu o\rho\phi\eta'$ (form)] if they have the same multiplication table.

Table 3 Multiplication table for the group composed of matrices given by Eqs. (2). The written order of the successive operations is specified by the arrow in the upper left-hand corner.

↗	E	A	B	C	D	F
E	E	A	B	C	D	F
A	A	B	E	F	C	D
B	B	E	A	D	F	C
C	C	D	F	E	A	B
D	D	F	C	B	E	A
F	F	C	D	A	B	E

8.3 PERMUTATIONS

A permutation of n ordered symbols is a reordering of the symbols. Thus if the symbols are initially numbered from 1 to n in order, a permutation results in an order $i_1, i_2, i_3, \ldots i_n$. This operation is often represented by

$$\hat{P} = \begin{bmatrix} 1 & 2 & 3 & \cdots & n \\ i_1 & i_2 & i_3 & \cdots & i_n \end{bmatrix}, \tag{3}$$

where the symbols are written in their initial order in the first row and in their final order in the second.* The identity operation is of course

$$\hat{P} = \begin{bmatrix} 1 & 2 & 3 & \cdots & n \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}.$$

The inverse of the operation given in Eq. (3) is simply

$$\hat{P}^{-1} = \begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}. \tag{4}$$

It is the operation which restores the original order.

The combination law ("product") is the result of two successive permutations, say, $\hat{P}_1$ and $\hat{P}_2$. If $\hat{P}_1$ operates on the initially ordered symbols, $\hat{P}_2$ then carries out the permutation of the order established by $\hat{P}_1$. As a simple example, consider three identical "objects" identified as $1, 2, 3, \ldots$. If

$$\hat{P}_1 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} \text{ and } \hat{P}_2 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}$$

---

*If no confusion results, the first row can be suppressed. Then, the general permutation operation is written simply as $\hat{P} = \begin{bmatrix} i_1 & i_2 & i_3 & \cdots & i_n \end{bmatrix}$.

are the two possible permutations; their product becomes

$$\hat{P}_2\hat{P}_1 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}. \tag{5}$$

In general, multiplication is not commutative, as can be demonstrated with this example, namely,

$$\hat{P}_1\hat{P}_2 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}. \tag{6}$$

As it can be shown that multiplication is associative, the permutation operations form a group. Furthermore, as shown in Chapter 10, there are $n!$ permutations of the n symbols. Hence, the corresponding group is of order $n!$, or six in this example.

It should be emphasized that in the above presentation of permutation operations, they were carried out on symbols, rather than physical objects. One symbol was exchanged for another as a result of a "paper operation". The application of this principle in physical systems must be made with care. When it is said that the "exchange of two identical particles yields the following results", it must be understood that it is the exchange of identity of the particles, such as labels or coordinates that has been made.

Imagine the following story – a sort of "gedanken experiment".* A pair of identical twins is born to Mama X. She places each of them in his crib, but to distinguish them she puts a sticker on each forehead. One twin is identified as Matthew and the other as Mark. Mama X leaves the room to let the two babies sleep. Then, big sister comes into the room and decides to play a joke. She exchanges the two stickers. When Mama X returns she comes to the immediate conclusion that the two infants have been exchanged! Assuming that the two babies are identical, she is perhaps right. However, nobody expended the energy to lift them and make the physical exchange of their positions. The operation was simply a permutation of the two symbols used to identify the infants. Such is a permutation operation, as defined above.

8.4 CONJUGATE ELEMENTS AND CLASSES

Two elements A and B of a group are said to be conjugate if

$$A = R^{-1}BR, \tag{7}$$

---

*This is the sort of experiment that is carried out with the application of thoughts (German: *Dänken*), rather than hands (French: *manipulation*).

where R is a member of the group. If Eq. (7) is premultiplied by R and postmultiplied by R^{-1}, the result is

$$B = RAR^{-1}. \tag{8}$$

As R^{-1} is also a member of the group, Eqs. (7) and (8) are symmetric and the elements A and B are mutually conjugate.

If P and Q are both conjugate to the same element W, they are mutually conjugate. This relation can be demonstrated as follows. Given $P = S^{-1}WS$ and $Q = TWT^{-1}$, where S and T are also members of the group,

$$W = T^{-1}QT. \tag{9}$$

Substitution of Eq. (9) in the expression for P leads to

$$P = S^{-1}(T^{-1}Q\,T)S = (T\,S)^{-1}Q(T\,S), \tag{10}$$

where the product TS and its inverse are also members of the group. Note that the relation $S^{-1}T^{-1} = (TS)^{-1}$, as given by Eq. (7-37), and the associative property of the elements have been employed in deriving Eq. (10).

The ensemble of elements that are mutually conjugate form a class. The concept of a class is most easily demonstrated by an example. The multiplication table for the group of matrices defined by Eq. (2) is given in Table 3. With its use the relations

$$\begin{aligned} E^{-1}AE &= A \\ A^{-1}AA &= A \\ B^{-1}AB &= A \\ C^{-1}AC &= B \\ D^{-1}AD &= B \\ F^{-1}AF &= B \end{aligned} \tag{11}$$

can be verified (problem 3). In this example the element A was arbitrarily chosen and the similarity transformations were carried out successively with the use of each element of the group. If this procedure is repeated with the choice of another element of the group, *e.g.*, $E^{-1}BE = B$, $A^{-1}BA = B \ldots$, it will be found that the results of the similarity transformations is to form three classes. They are composed of the elements: E, which is always in a class by itself, the two elements A and B, which are "mixed" as shown above, and a third composed of the operations C, D and F (problem 4).

8.5 MOLECULAR SYMMETRY

Although all molecules are in constant thermal motion, when all of their atoms are at their equilibrium positions, a specific geometrical structure can usually be assigned to a given molecule. In this sense these molecules are said to be "rigid".* The first step in the analysis of the structure of a molecule is the determination of the group of operations that characterizes its symmetry. Each symmetry operation (aside from the trivial one, E) is associated with an element of symmetry. Thus for example, certain molecules are said to be planar. Well known examples are water, boron trifluoride and benzene, whose structures can be drawn on paper in the forms shown in Fig. 1.

The symmetry operations and their descriptions, as well as the notation established by Schönflies, are given in Table 4. To identify the symmetry group it is necessary to visualize all of the operations which characterize the structure of the molecule. As there is always at least one point in space that is not affected by any of the symmetry operations, the group is referred to as a point group.† It is sometimes useful to employ a systematic method to determine the point group (see Appendix VI), although with experience and the aid of molecular models the operations are usually not difficult to find (try problems 5–8).

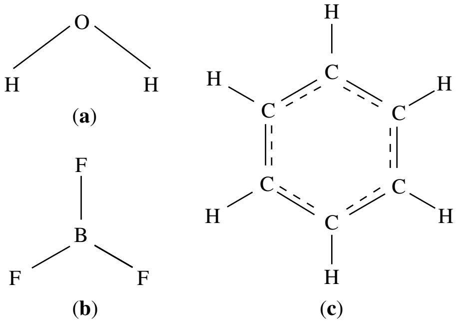


Fig. 1 Planar Molecules: (**a**) water, (**b**) boron trifluoride, (**c**) benzene.

---

*Some molecules are not "rigid" as they contain functional groups that have a certain orientational freedom. For example, dimethyl acetylene, $CH_3{-}C{\equiv}C{-}CH_3$, in which the two methyl groups rotate with little hindrance, cannot be described by a specific geometry.

†In crystallography additional operations involving translation of the unit cell are employed. The resulting groups are known as space groups (see Section 8.13).

Table 4 Symmetry operations used to describe the point symmetry of a "rigid" molecule.

Symmetry operation	Symbol	Symmetry element
Identity	E	—
Reflection in a plane	σ	Plane of symmetry
Rotation k times by $2\pi/n$	C_n^k	Proper rotation axis of order n
Rotation by $2\pi/n$ followed by a reflection in the plane $\perp$ to the axis	S_n^k	Improper rotation axis of order n
Inversion with respect to a center	$i \equiv S_2^1$	Center of symmetry

The molecules shown in Fig. 1 are planar; thus, the paper on which they are drawn is an element of symmetry and the reflection of all points through the plane yields an equivalent (congruent) structure. The process of carrying out the reflection is referred to as the symmetry operation σ. However, as the atoms of these molecules are essentially point masses, the reflection operations are in each case simply the inversion of the coordinate perpendicular to the plane of symmetry. Following certain conventions, the reflection operation corresponds to $z \to -z$ for BF_3 and benzene, as it is the z axis that is chosen perpendicular to the plane, while it is $x \to -x$ for water. It should be evident that the symmetry operation has an effect on the chosen coordinate systems, but not on the molecule itself.

In general, a molecule may have a number of elements of symmetry. For example, the water molecule has, in addition to the plane of symmetry indicated above, a second plane of symmetry perpendicular to the first one (see Fig. 2). The latter plane (z, x) is a consequence of the geometrical equivalence of the two hydrogen atoms in this molecule. The symmetry operation of reflection through the plane z, x, usually represented by the symbol $\sigma(zx)$, results in the permutation (exchange) of the coordinates of the two hydrogen atoms, the permutation [2,1]. It also inverts the y axis; that is, $y \to -y$. It should be noted that these inversions (of one axis in each case) result in the change from a right-handed coordinate system to a left-handed one (see Section 4.4). The same result is obtained by the inversion of all three axes. However, it should not be forgotten that the molecule is not "aware" of the choice of coordinate system, nor of the paper operations that may have been made.

Finally, continuing with the example of the water molecule, the operation C_2 is also characteristic of its symmetry. This operation results in the inversions $x \to -x$ and $y \to -y$, as well as the permutation [2,1]. Because two

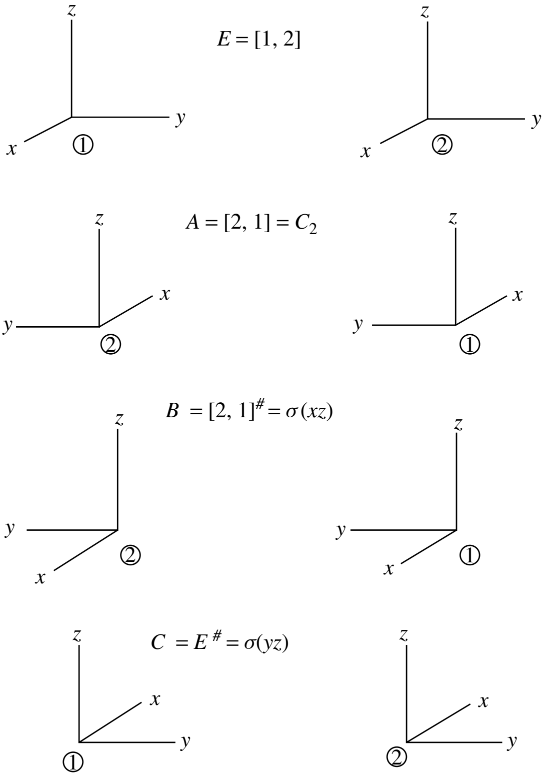


Fig. 2 The effect of the symmetry operations on the Cartesian displacement coordinates of the two hydrogen atoms in the water molecule. The sharp (#) indicates the inversion of a coordinate axis, resulting in a change in "handedness" of the Cartesian coordinate system.

inversions have been made, the coordinate system remains right-handed. Obviously, the effects of the symmetry operations on the coordinates of the oxygen atom are only those of the coordinate inversions, as the oxygen atom has no symmetrically equivalent partner.

The four operations which form the symmetry group for the water molecule are represented in Fig. 2. It can be easily verified that the multiplication table for these symmetry operations is that already developed (Table 2). Thus, the symmetry group of the water molecule is isomorphic with the four-group.

The operations shown in Fig. 2. can be easily recast in matrix language. For completeness, the displacement coordinates of the oxygen atom (numbered ③) will be included. The Cartesian components of the displacement of each atom from its equilibrium position are arranged as a vector,

$$\xi = \left(\begin{array}{c} \Delta x_1 \\ \Delta y_1 \\ \Delta z_1 \\ \hdashline \Delta x_2 \\ \Delta y_2 \\ \Delta z_2 \\ \hdashline \Delta x_3 \\ \Delta y_3 \\ \Delta z_3 \end{array}\right) \qquad (12)$$

The result of a given symmetry operation, $\hat{R}$, is to produce the vector ξ', given by

$$\xi' = \boldsymbol{R}\xi, \qquad (13)$$

where $\boldsymbol{R}$ is the matrix that represents the operation $\hat{R}$. Matrix representations can be constructed as follows:

$$\boldsymbol{E} = \left(\begin{array}{ccc:ccc:ccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right), \qquad (14)$$

$$\boldsymbol{A} = \left(\begin{array}{ccc:ccc:ccc} 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hdashline -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right), \qquad (15)$$

$$\boldsymbol{B} = \left(\begin{array}{ccc|ccc|ccc}
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\hline
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right), \tag{16}$$

$$\boldsymbol{C} = \left(\begin{array}{ccc|ccc|ccc}
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right). \tag{17}$$

The nine-by-nine matrix that represents a given operation is composed of three-by-three submatrices. Those submatrices which lie along the principal diagonal contain nonzero elements if the coordinates of the atom in question are not permuted by the operation, *e.g.* E and C in this example Otherwise, nonzero elements appear in the appropriate off-diagonal submatrices, as indicated in the representations of A and B [Eqs. (15) and (16)]. The elements -1 and 1 in the various submatrices indicate, respectively, inversion or not of the coordinate involved.

Although a certain amount of effort is involved, it can be demonstrated that the symmetry operations represented by the above matrices [Eqs. (14) to (17)] form the four-group, as defined by Table 2 (problem 9). In the Schönflies notation this group is known as $\mathscr{C}_{2\mathrm{v}}$. The fact that the labels ① and ② can be permuted establishes the symmetrical equivalence of the two hydrogen atoms in the water molecule. Furthermore, the particular operations, E and C, which do not involve the permutation, form a group of order two ($\mathscr{C}_{\mathrm{s}}$), a subgroup of the four-group.

As a second example of molecular symmetry, consider the ammonia molecule. It has three symmetrically equivalent hydrogen atoms, but it is not planar. The symmetry can thus be described by all possible permutations of the labels on the three hydrogen atoms. The order of the group is then equal to $3! = 6$, formed by the permutations:

$$\left.\begin{bmatrix}1 & 2 & 3\end{bmatrix}=E\right\}\qquad\qquad \text{Class } E$$

$$\left.\begin{aligned}\begin{bmatrix}3 & 1 & 2\end{bmatrix}&=C_3^1=A\\ \begin{bmatrix}2 & 3 & 1\end{bmatrix}&=C_3^2=B\end{aligned}\right\}\qquad\qquad \text{Class } 2C_3$$

$$\left.\begin{aligned}\begin{bmatrix}1 & 3 & 2\end{bmatrix}^{\#}&=\sigma_1=C\\ \begin{bmatrix}3 & 2 & 1\end{bmatrix}^{\#}&=\sigma_2=D\\ \begin{bmatrix}2 & 1 & 3\end{bmatrix}^{\#}&=\sigma_3=F\end{aligned}\right\}\qquad\qquad \text{Class } 3\sigma_{\mathrm{v}}. \tag{18}$$

The symbols $A, \ldots, F$ refer to the multiplication table (Table 3), as can be verified (see problem 10). The notation for the classes given in Eqs. (18) is again that of Schönflies.

With the use of the Cartesian displacement coordinates defined in Fig. 3, a basis vector is of the form

$$\xi=\begin{pmatrix}\Delta x_1\\ \Delta y_1\\ \Delta z_1\\ \text{- - - -}\\ \Delta x_2\\ \Delta y_2\\ \Delta z_2\\ \text{- - - -}\\ \Delta x_3\\ \Delta y_3\\ \Delta z_3\\ \text{- - - -}\\ \Delta x_4\\ \Delta y_4\\ \Delta z_4\end{pmatrix}. \tag{19}$$

With some patience it is found that the operation $\begin{bmatrix}3 & 1 & 2\end{bmatrix}=C_3^1$ can be represented by the matrix

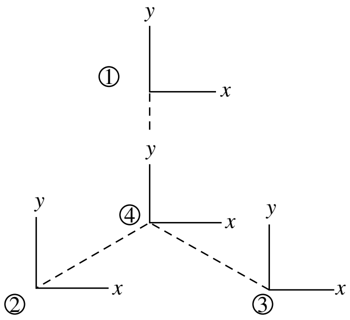


Fig. 3 Cartesian displacement coordinates for the ammonia molecule. The $z(C_3)$ axis is perpendicular to the plane of the paper (which is not a plane of symmetry).

$$
\boldsymbol{A} = \left(\begin{array}{ccc:ccc:ccc:ccc}
0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\hdashline
-\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right). \tag{20}
$$

It should be noted that each nonzero submatrix is of the form given by Eq. (7-104), with $\varphi = 120°$. The corresponding representation for the operation B is readily obtained with $\varphi = 240°$.

Similarly, with respect to the same basis the representation of the operation $\begin{bmatrix} 1 & 3 & 2 \end{bmatrix}^{\#} = \sigma_1$ is of the relatively simple form

$$\boldsymbol{C} = \left(\begin{array}{ccc:ccc:ccc:ccc}
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right), \tag{21}$$

while the matrix which represents $\begin{bmatrix} 3 & 2 & 1 \end{bmatrix}^{\#} = \sigma_2$ is

$$\boldsymbol{D} = \left(\begin{array}{ccc:ccc:ccc:ccc}
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right). \tag{22}$$

The representation of $\begin{bmatrix} 2 & 1 & 3 \end{bmatrix}^{\#} = \sigma_3$ is analogous (problem 12).

8.6 THE CHARACTER

In the examples presented in the previous section, the vectors ξ of displacement coordinates [Eqs. (12) and (19)] were used as a basis. It should not be surprising that the matrices employed to represent the symmetry operations have different forms depending on the basis coordinates. In effect, there is an infinite number of matrices that can serve as representations of a given symmetry operation. Nevertheless, there is one quantity that is characteristic of the operation – the trace of the matrix – as it is invariant under a change of basis coordinates. In group theory it is known as the character.

To show that the character is invariant under a change of basis coordinates, consider the effect of an operation $\hat{R}$ on the basis vector ξ; thus,

$$\xi' = \boldsymbol{R}\xi, \tag{23}$$

as given for example by Eq. (13). Now choose another basis $\boldsymbol{\eta}$, which is linearly related to the original one, namely

$$\boldsymbol{\eta} = \boldsymbol{S}\xi. \tag{24}$$

The same transformation can be applied to the result given by Eq. (23). Then,

$$\boldsymbol{\eta}' = \boldsymbol{S}\xi', \tag{25}$$

and, from Eq. (25) $\xi = \boldsymbol{S}^{-1}\boldsymbol{\eta}$. Substitution in Eq. (23), with the use of Eq. (25), yields

$$\boldsymbol{\eta}' = \boldsymbol{S}\xi' = \boldsymbol{S}\ \boldsymbol{R}\xi = (\boldsymbol{SRS}^{-1})\boldsymbol{\eta}. \tag{26}$$

This result shows that the matrix $(\boldsymbol{SRS}^{-1})$ in the new basis corresponds to $\boldsymbol{R}$ the original one. The relation between them is a similarity transformation (see Section 7.10). It is now necessary to demonstrate that the character of a matrix transformation is invariant under a similarity transformation.

In Section 7.3 it was shown that the trace of the product of two matrices is independent of the order of multiplication. Thus, $Tr(\boldsymbol{AB}) = Tr(\boldsymbol{BA})$, a result that can be applied to the similarity transformation of Eq. (26), *e.g.*

$$\begin{aligned} Tr(\boldsymbol{SRS}^{-1}) &= Tr[\boldsymbol{S}(\boldsymbol{RS}^{-1})] = Tr[(\boldsymbol{RS}^{-1})\boldsymbol{S}] \\ &= Tr[\boldsymbol{R}(\boldsymbol{S}^{-1}\boldsymbol{S})] = Tr(\boldsymbol{R}), \end{aligned} \tag{27}$$

where the associative property of the matrices has been applied. Here it is shown, not only that the character of a matrix transformation is invariant under a similarity transformation, but also that all of the representations of the same class have the same character. Both of these results are fundamental in the following applications of group theory.

8.7 IRREDUCIBLE REPRESENTATIONS

As indicated above there may be many equivalent matrix representations for a given operation in a point group. Although the form depends on the choice of basis coordinates, the character is independent of such a choice. However, for each application there exists a particular set of basis coordinates in terms of which the representation matrix is reduced to block-diagonal form. This result is shown symbolically in Fig. 4. It can be expressed mathematically by the relation

$$\mathbf{\Gamma} = \sum_{\gamma} n^{(\gamma)} \mathbf{\Gamma}^{(\gamma)}, \tag{28}$$

where $\mathbf{\Gamma}$ is a given (reducible) representation and Fig. 4 shows in schematic form the result of reducing it. The matrices appearing along the principal diagonal of $\mathbf{\Gamma}$ are referred to as the irreducible representations, because there is no transformation that can further simplify them.

The reduction of a representation is accomplished with the application of a similarity transformation. Thus, there exists a matrix $\boldsymbol{D}$ such that the product $\boldsymbol{D}^{-1}\mathbf{\Gamma}\,\boldsymbol{D}$ reduces the representation $\mathbf{\Gamma}$ to block diagonal form. Not only is the character preserved under this transformation, but the resulting reduced representations also belong to the symmetry group in question. In most applications it is not necessary to find the matrix $\boldsymbol{D}$, nor is it in general important to calculate the irreducible representations. It is usually sufficient to find their diagonal elements, or at the least, their characters. This problem will be illustrated in the following section.

It should be apparent that the summation indicated in Eq. (28) is special. It indicates that the direct sum is to be taken; that is, the various irreducible representations are arranged in arbitrary order along the diagonal of the reduced

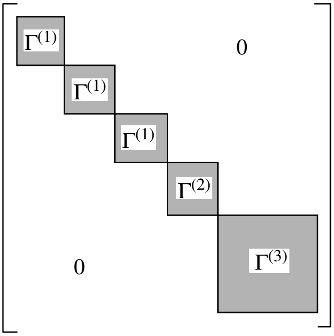


Fig. 4 Block-diagonal form of a representation matrix: The reduced representation.

representation. Each irreducible representation, identified by the superscript (γ) may appear more than once, or not at all. Thus, $n^{(\gamma)}$ is equal to the number of times that the irreducible representation $\mathbf{\Gamma}^{(\gamma)}$ appears in the reduced representation.

Before going on to consider applications of group theory in physical problems, it is necessary to discuss several general properties of irreducible representations. First, suppose that a given group $\mathcal{G}$ is of order g and that the g operations have been collected into k different classes of mutually conjugate operations. It can be shown that the group $\mathcal{G}$ possesses precisely k nonequivalent irreducible representations, $\mathbf{\Gamma}^{(1)}$, $\mathbf{\Gamma}^{(2)}, \ldots, \mathbf{\Gamma}^{(k)}$, whose dimensions $d_1, d_2, \ldots, d_k$ satisfy the relation

$$d_1^2 + d_2^2 + \ldots + d_k^2 = g. \tag{29}$$

This statement is often taken as a basic theorem of representation theory. It is found that for any symmetry group there is only one set of k integers (zero or positive), the sum of whose squares is equal to g, the order of the group. Hence, from Eq. (29), $n^{(\gamma)}$, the number of times that each irreducible representation appears in the reduced representation, as well as its dimension, can be determined for any group.

It was shown earlier that all operations belonging to the same class have the same characters. The classes, which are denoted by $\mathscr{K}_1, \mathscr{K}_2, \ldots, \mathscr{K}_k$, are composed of $g_1, g_2, \ldots, g_k$ operations, respectively. Finally the symbol $\chi_j^{(\gamma)}$ is used for the character of the γth representation of an operation of class $\mathscr{K}_j$.

The above definitions allow the property of orthogonality of the characters to be stated in the form

$$\frac{1}{g}\sum_{j=1}^{k} g_j \chi_j^{(\gamma')\star} \chi_j^{(\gamma)} = \delta_{\gamma',\gamma}, \tag{30}$$

where $\chi_j^{(\gamma')\star}$ is the complex conjugate of the character $\chi_j^{(\gamma')}$ and $\delta_{\gamma',\gamma}$ is the Kronecker delta. The orthogonality of the characters will be used repeatedly in the group-theoretical applications which follow.* Because of the property expressed by Eq. (27), each irreducible representation can be specified by its set of characters. These quantities are conveniently listed for each group in a character table.

---

*Equation (30) can be derived from the general orthogonality theorem, Eq. (44).

Table 5 The general form of a character table.

$\mathcal{G}$	$\mathscr{K}_1 = E$	$\mathscr{K}_2$	$\mathscr{K}_3$	$\cdots$	$\mathscr{K}_k$
$\mathbf{\Gamma}^{(1)}$	$\chi_1^{(1)}$	$\chi_2^{(1)}$	$\chi_3^{(1)}$	$\cdots$	$\chi_k^{(1)}$
$\mathbf{\Gamma}^{(2)}$	$\chi_1^{(2)}$				
$\mathbf{\Gamma}^{(3)}$	$\chi_1^{(3)}$				
$\vdots$	$\vdots$				
$\mathbf{\Gamma}^{(k)}$	$\chi_1^{(k)}$	$\cdots$	$\cdots$	$\cdots$	$\chi_k^{(k)}$

8.8 CHARACTER TABLES

The tables of characters have the general form shown in Table 5. Each column represents a class of symmetry operation, while the rows designate the different irreducible representations. The entries in the table are simply the characters (traces) of the corresponding matrices. Two specific properties of the character tables will now be considered.

As shown earlier, the operation E is always in a class by itself, as it commutes with all other operations of the group. It is identified with $\mathscr{K}_1$, the arbitrarily chosen first class of operation. In a given representation the operation E corresponds to a unit matrix whose order is equal to the dimension of the representation. Hence, the resulting character, the sums of the diagonal elements, is also equal to the dimension of the representation. The dimension of each representation can thus be easily determined by inspection of the corresponding entry in the first column of characters in the table.

A further property of the character tables arises from the fact that every symmetry group has an irreducible representation that is invariant under all of the group operations. This irreducible representation is a one-by-one unit matrix (the number one) for every class of operation. Obviously, the characters, are all then equal to one. As this irreducible representation is by convention taken to be $\mathbf{\Gamma}^{(1)}$, the first row of all character tables consists solely of ones. The significance of the character tables will become more apparent by consideration of an example.

The operations defined by Eqs. (18) form the group of order $n! = 6$, which in the Schönflies notation is known as $\mathscr{C}_{3\text{v}}$. As the operations are divided into $k = 3$ classes, Eq. (29) becomes

$$d_1^2 + d_2^2 + d_3^2 = 6. \tag{31}$$

Only the set of integers 1, 1, 2 satisfies this relation, the order being arbitrary. In this group there are two different irreducible representations of order one and one of order two. Thus, the characters appearing in the column headed

Table 6 The character table for the group $\mathcal{C}_{3\mathrm{v}}$.

$\mathcal{C}_{3\mathrm{v}}$	$[1\ 2\ 3] = E$	$\left.\begin{matrix}[3\ 1\ 2]\\ [2\ 3\ 1]\end{matrix}\right\} = 2C_3$	$\left.\begin{matrix}[1\ 3\ 2]^{\#}\\ [3\ 2\ 1]^{\#}\\ [2\ 1\ 3]^{\#}\end{matrix}\right\} = 3\sigma_v$
$\boldsymbol{\Gamma}^{(1)} \equiv \mathrm{A}_1$	1	1	1
$\boldsymbol{\Gamma}^{(2)} \equiv \mathrm{A}_2$	1	1	-1
$\boldsymbol{\Gamma}^{(3)} \equiv \mathrm{E}$	2	-1	0

E (for the identity) in Table 6 are accounted for. Furthermore, the totally symmetric representation is $\boldsymbol{\Gamma}^{(1)} \equiv \mathrm{A}_1$; the latter notation is that usually used by spectroscopists.* The construction of the remainder of the character table is accomplished by application of the orthogonality property of the characters [see Eq. (30) and problem 13]. Standard character tables have been derived in this way for the more common groups, as given in Appendix VIII.

In molecular spectroscopy the term symmetry species is usually substituted for the rather unwieldy expression "irreducible representation".† Furthermore, the word degeneracy, as employed in spectroscopy, is often used to specify their dimension.

The group developed above to describe the symmetry of the ammonia molecule consisted only of the permutation operations. However, if the triangular pyramid corresponding to this structure is flattened, it becomes planer in the limit. The BF_3 molecule shown in Fig. 1b is an example of this symmetry. In this case it becomes possible to invert the coordinate perpendicular to the plane of the molecule, the z axis. Obviously, the operation of reflection in the (horizontal) plane of the molecule, σ_{h}, is identical. It is easy, then, to identify the irreducible representations A' and A'' as symmetric or antisymmetric, respectively, under the coordinate inversion. The group composed of the identity and the inversion of the z axis is then $\mathcal{C}_{1\mathrm{h}} \equiv \mathcal{C}_{\mathrm{s}}$, whose character table is of the form of Table 7.

Table 7 The character table for the group $\mathcal{C}_{1\mathrm{h}} \equiv \mathcal{C}_{\mathrm{s}}$.

$\mathcal{C}_{1\mathrm{h}} \equiv \mathcal{C}_{\mathrm{s}}$	E	$E^{\#} = \sigma_{\mathrm{h}}$
A'	1	1
A''	1	-1

---

*The symbols A_1, A_2, E are sometimes replaced by a_1, a_2, e, respectively.

†Note, however, that the word "class" should not be used for the symmetry species.

Table 8 The character table for the group $\mathscr{D}_{3\mathrm{h}}$.

$\mathscr{D}_{3\mathrm{h}}$	E	$\left.\begin{matrix}[3\ 1\ 2]\\ [2\ 3\ 1]\end{matrix}\right\} = 2C_3$	$\left.\begin{matrix}[1\ 3\ 2]\\ [3\ 2\ 1]\\ [2\ 1\ 3]\end{matrix}\right\} = 3C_2$	$E^{\#} = \sigma_{\mathrm{h}}$	$\left.\begin{matrix}[3\ 1\ 2]^{\#}\\ [2\ 3\ 1]^{\#}\end{matrix}\right\} = 2S_3$	$\left.\begin{matrix}[1\ 3\ 2]^{\#}\\ [3\ 2\ 1]^{\#}\\ [2\ 1\ 3]^{\#}\end{matrix}\right\} = 3\sigma_v$
A_1'	1	1	1	1	1	1
A_2'	1	1	−1	1	1	−1
E'	2	−1	0	2	−1	0
A_1''	1	1	1	−1	−1	−1
A_2''	1	1	−1	−1	−1	1
E''	2	−1	0	−2	1	0

Because the operations of the group $\mathscr{C}_{1\mathrm{h}}$ commute with all of the operations of the group $\mathscr{C}_{3\mathrm{v}}$, a group of order 12 can be formed from all possible products of the operations of these two groups. The product group thus formed is called $\mathscr{D}_{3\mathrm{h}}$. It is not surprising that its irreducible representations are found by taking the direct product of those of the two smaller groups. The character table given as Table 8 can be calculated from the definition of the direct product of two matrices, $\mathscr{C}_{3\mathrm{v}} \otimes \mathscr{C}_{1\mathrm{h}} = \mathscr{D}_{3\mathrm{h}}$ (see Section 7.3).

The character table for $\mathscr{D}_{3\mathrm{h}}$ is seen to be composed of four submatrices, three of which are just the character table of $\mathscr{C}_{3\mathrm{v}}$ (Table 6). In forming the submatrix in the lower, right position the elements have been multiplied by -1, the character of A$''$ under the operation $E^{\#} = \sigma_{\mathrm{h}}$. The direct product method, as employed here, is used to calculate a number of the more complicated character tables.

8.9 REDUCTION OF A REPRESENTATION: THE "MAGIC FORMULA"

It was stated above that the number of irreducible representations in any group is equal to the number of classes, k. Thus, Eq. (28) becomes

$$\mathbf{\Gamma} = \sum_{\gamma=1}^{k} n^{(\gamma)} \mathbf{\Gamma}^{(\gamma)}. \tag{32}$$

Here, $\mathbf{\Gamma}$, as given by the direct sum, is a (reducible) representation of a given operation, $\hat{R}$. Its trace is the character, a quantity that is independent of the choice of basis coordinates. As χ_R is merely the sum of the diagonal elements of $\mathbf{\Gamma}$, it is also equal to the sum of the traces of the individual submatrices

$\boldsymbol{\Gamma}^{(\gamma)}$, each multiplied by $n^{(\gamma)}$, the number of times that each $\boldsymbol{\Gamma}^{(\gamma)}$ appears along the diagonal of $\boldsymbol{\Gamma}$. It follows that

$$\chi_R = \sum_{\gamma=1}^{k} n^{(\gamma)} \chi_R^{(\gamma)}, \tag{33}$$

an expression that holds for each operation of the group. However, as all operations of a given class have the same characters, only k different equations of this type need to be formed, one for each class. Hence,

$$\chi_j = \sum_{\gamma=1}^{k} n^{(\gamma)} \chi_j^{(\gamma)}, \quad j = 1, 2, \ldots, k \tag{34}$$

where $\chi_j^{(\gamma)}$ is the character of the irreducible representation $\boldsymbol{\Gamma}^{(\gamma)}$ of an operation of class $\mathscr{K}_j$.

This set of simultaneous equations [Eq. (34)] can be solved with aid of the orthogonality properties of the characters, as given by Eq. (30). Multiplication of each side of Eq. (34) by $\chi_j^{(\gamma')\star}$, weighted by the factor g_j and summed over the class index j, yields the expression

$$\begin{aligned}\sum_{j=1}^{k} g_j \chi_j^{(\gamma')\star} \chi_j &= \sum_{j=1}^{k} g_j \chi_j^{(\gamma')\star} \left[\sum_{\gamma=1}^{k} n^{(\gamma)} \chi_j^{(\gamma)} \right] \\ &= \sum_{\gamma=1}^{k} n^{(\gamma)} \left[\sum_{j=1}^{k} g_j \chi_j^{(\gamma')\star} \chi_j^{(\gamma)} \right].\end{aligned} \tag{35}$$

However, from Eq. (30) the quantity in brackets is equal to $g\delta_{\gamma,\gamma'}$ and Eq. (35) becomes

$$\sum_{j=1}^{k} g_j \chi_j^{(\gamma')\star} \chi_j^{(\gamma)} = \sum_{\gamma=1}^{k} n^{(\gamma)} g \delta_{\gamma',\gamma} = g n^{(\gamma)} \tag{36}$$

or

$$n^{(\gamma)} = \frac{1}{g} \sum_{j=1}^{k} g_j \chi_j^{(\gamma)\star} \chi_j. \tag{37}$$

This result (problem 14) allows the coefficients $n^{(\gamma)}$ to be calculated. Thus, with a knowledge of the symmetry group and the corresponding table of characters, the structure of the reduced representation can be determined. Equation (37) is of such widespread applicability that it is referred to by many students of group theory as the “magic formula”.

To illustrate the application of Eq. (37), consider the ammonia molecule with the system of 12 Cartesian displacement coordinates given by Eq. (19) as the basis. The reducible representation for the identity operation then corresponds to the unit matrix of order 12, whose character is obviously equal to 12. The symmetry operation $A = C_3^1$ of Eq. (18) is represented by the matrix of Eq. (20) whose character is equal to zero. The same result is of course obtained for the operation B, as it belongs to the same class. For the class $3\sigma_{\mathrm{v}}$ the character is equal to two, as exemplified by the matrices given by Eqs. (21) and (22) for the operations C and D, respectively. The representation of the operation F is analogous to D (problem 12).

The characters of the reducible representations for this example are employed to calculate the structure of the reduced representation. As pointed out in Section 7.12, the character is equal to $2\,cos\,\varphi \pm 1$ for a given geometrical operation. Thus, if there are m_j coordinate systems ("atoms") that are not permuted by a symmetry operation of a given class j, the character of the reducible representation is given by $\chi_j = m_j(2\,cos\,\varphi_j \pm 1)$. The character table (Table 6) can now be rewritten with the addition of these results in the last row. Application of the magic formula yields the values of $n^{(\gamma)}$ given in the last column of Table 9 (problem 15). Thus, the structure of the reduced representation is obtained without any specific knowledge of the irreducible representations. It should be noted that the sum of the numbers in the last column, weighted by the degeneracies given in the column headed E, yields $3 + 1 + 2 \times 4 = 12$. It is necessary (but not sufficient) that this value be equal to the order of the reducible representations, the last value in the column headed E.

Table 9 The symmetry group $\mathcal{C}_{3\mathrm{v}}$ and the reduction of the representations of the Cartesian displacements for the ammonia molecule.

$\mathcal{C}_{3\mathrm{v}}$	E	$2C_3$	$3\sigma_{\mathrm{v}}$	$n^{(\gamma)}$
$\mathbf{\Gamma}^{(1)} \equiv \mathrm{A}_1$	1	1	1	3
$\mathbf{\Gamma}^{(2)} \equiv \mathrm{A}_2$	1	1	-1	1
$\mathbf{\Gamma}^{(3)} \equiv \mathrm{E}$	2	-1	0	4
χ_j	12	0	2	

8.10 THE DIRECT PRODUCT REPRESENTATION

The direct product of two matrices was defined in the previous chapter by Eq. (7-30), *viz.*

$$[\boldsymbol{A} \otimes \boldsymbol{B}]_{ik,jl} = a_{ij} b_{kl}. \tag{38}$$

Thus, for example, if

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ and } \boldsymbol{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

their direct product is given by

$$\boldsymbol{A} \otimes \boldsymbol{B} = \boldsymbol{C} = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix}. \tag{39}$$

Clearly,

$$\begin{aligned}(Tr\,\boldsymbol{A})(Tr\,\boldsymbol{B}) &= (a_{11} + a_{22})(b_{11} + b_{22}) \\ &= a_{11}b_{11} + a_{11}b_{22} + a_{22}b_{11} + a_{22}b_{22} = Tr\boldsymbol{C}.\end{aligned} \tag{40}$$

This general principle is applicable to representations, where the term character is used to describe the trace (Section 8.6).

To illustrate the importance of the direct product representation, consider the very simple classification of functions (in one dimension) as odd or even. In the context of group theory these properties correspond to the effect of the inversion of the coordinate, say $x \to -x$. Thus, even or odd functions belong to the irreducible representations $\mathrm{A_g}$ or $\mathrm{A_u}$, respectively. The subscripts are g (Gerade) and u (Ungerade), corresponding to even and odd under the inversion operation. In this example the usual rules, $\mathrm{g \times g = g}$, $\mathrm{u \times u = g}$ and $\mathrm{g \times u = u \times g = u}$ are expressed as $\mathrm{A_g \otimes A_g = A_g}$, $\mathrm{A_u \otimes A_u = A_g}$ and $\mathrm{A_g \otimes A_u = A_u \otimes A_g = A_u}$. These results should be obvious from the character table for the group $\mathcal{C}_i$, given as Table 10.

As indicated in Section 3.4, the integral of an odd function, taken between symmetric limits, is equal to zero. More generally, the integral of a function that is not symmetric with respect to all operations of the appropriate point group will vanish. Thus, if the integrand is composed of a product of functions, each of which belongs to a particular irreducible representation, the overall symmetry is given by the direct product of these irreducible representations.

Table 10 Character table for the group $\mathcal{C}_i$.

$\mathcal{C}_i$	E	$\boldsymbol{i}$
$\mathrm{A_g}$	1	1
$\mathrm{A_u}$	1	−1

Although the resulting direct product may not be reduced, it can be made so by application of the magic formula, or often by inspection. The nonvanishing of the integral is then determined by the existence of the totally symmetric representation in the resulting direct sum. This procedure will be illustrated by the development of spectroscopic selection rules in Section 12.3.3.

In the following chapter this brief outline of representation theory will be applied to several problems in physical chemistry. It is first necessary, however, to show how functions can be adapted to conform to the natural symmetry of a given problem. It will be demonstrated that this concept is of particular importance in the analysis of molecular vibrations and in the theory of molecular orbitals, among others. The reader is warned, however, that a serious development of this subject is above the level of this book. Hence, in the following section certain principles will be presented without proof.

8.11 SYMMETRY-ADAPTED FUNCTIONS: PROJECTION OPERATORS

Consider a vector $\boldsymbol{\phi}^{(\gamma)}$ composed of the set of orthonormal functions, $\phi_1^{(\gamma)}$, $\phi_2^{(\gamma)}, \ldots, \phi_{d_\gamma}^{(\gamma)}$ which forms a basis for the irreducible representation $\boldsymbol{\Gamma}^{(\gamma)}$. The result of a group operation $\hat{R}$ on a given member of the set, $\phi_t^{(\gamma)}$, is to produce a linear combination of the members of the set, as given by

$$\hat{R}\phi_t^{(\gamma)} = \boldsymbol{\Gamma}_t^{(\gamma)}(R)\boldsymbol{\phi}^{(\gamma)}. \tag{41}$$

The coefficients are the elements in row t of the irreducible representation.

It is instructive to write out Eq. (41) as

$$\hat{R}\phi_t^{(\gamma)} = \sum_s \Gamma_{s,t}^{(\gamma)}(R)\phi_{\mathrm{s}}^{(\gamma)}. \tag{42}$$

Multiplication of this expression by $\Gamma_{s',t'}^{(\gamma')\star}(R)$ and summation over the operations of the group yields

$$\sum_R \Gamma_{s',t'}^{(\gamma')\star}(R)\hat{R}\phi_t^{(\gamma)} = \sum_s \phi_s^{(\gamma)} \sum_R \Gamma_{s',t'}^{(\gamma')\star}(R)\Gamma_{s,t}^{(\gamma)}(R), \tag{43}$$

where the order of summation has been changed and the functions $\phi_s^{(\gamma)}$ have been removed from the sum over R. It can be shown from the general orthogonality properties of the irreducible representations that the second summation in Eq. (43) is equal to

$$\sum_R \Gamma_{s',t'}^{(\gamma')\star}(R)\Gamma_{s,t}^{(\gamma)}(R) = \frac{g}{d_\gamma}\delta_{\gamma,\gamma'}\delta_{s,s'}\delta_{t,t'}, \tag{44}$$

where g is the order of the group. This result is known as "the great orthogonality theorem", a special case of which has already been used [Eq. (30)] in deriving Eq. (37). Equation (43) then suggests that an operator be defined by

$$\hat{\mathcal{P}}^{(\gamma')}_{t',t'} = \frac{d_{\gamma'}}{g} \sum_{R} \Gamma^{(\gamma')\star}_{t',t'}(R)\hat{R}. \tag{45}$$

Then, Eq. (44) can be written as

$$\hat{\mathcal{P}}^{(\gamma')}_{s',t'} \phi^{(\gamma)}_{t} = \phi^{(\gamma)\star}_{s'} \delta_{t,t'}\delta_{\gamma,\gamma'}, \tag{46}$$

where the operator $\hat{\mathcal{P}}^{(\gamma')}_{s',t'}$ is known as a projection operator. It is sufficient to choose $t = t'$ so that only the diagonal elements of each $\boldsymbol{\Gamma}^{(\gamma')}$ are involved in Eq. (45). If the resulting operator, $\hat{\mathcal{P}}^{(\gamma')}_{t',t'}$ is applied to an arbitrary function $\phi^{(\gamma)}_{t}$, it will select a particular function $\phi^{(\gamma')}_{t}$ and set all others equal to zero. It is said, then, that this operator "projects" the particular function $\phi^{(\gamma)}_{t}$, or any component of it with the proper symmetry, and eliminates the others.

As a simple example, consider the operations E, A, B, and C, illustrated in Fig. 2. As this group is Abelian (see Table 2), all irreducible representations are of order one, *e.g.* $d_\gamma = 1$ for all γ. Furthermore, each irreducible representation is then just a number, the character, $\chi^{(\gamma)}_{j}$. The character table is relatively easy to construct with the aid of the method outlined above. The result is shown as Table 11.*

Now consider the arbitrary function $\phi = xy + zx + z$. The effect on this function of each operation $\hat{R}$ of the group is given in the last column of Table 11. This result is obvious if the symmetry operations are expressed in terms of coordinate inversions, as in Fig. 2. With $g = 4$ and $\Gamma^{(\gamma')\star}_{t,t}(R) = \chi^{(\gamma')}_{R}$, Eqs. (45) and (46) show that the products xy and zx "belong" to the irreducible representations A_2 and B_1, respectively, while the coordinate z is

Table 11 The character table for the group $\mathcal{C}_{2\mathrm{v}}$.

$\mathcal{C}_{2\mathrm{v}}$	E	$A = [2, 1] = C_2$	$B = [2, 1]^{\#} = \sigma(xz)$	$C = E^{\#} = \sigma(yz)$	Coordinates
$\Gamma_1 = \mathrm{A}_1$	1	1	1	1	z, x^2, y^2, z^2
$\Gamma_2 = \mathrm{A}_2$	1	1	−1	−1	xy
$\Gamma_3 = \mathrm{B}_1$	1	−1	1	−1	x, zx
$\Gamma_4 = \mathrm{B}_2$	1	−1	−1	1	y, yz

---

*Certain conventions are respected in the presentation of character tables, although the order of neither the classes (operations in this case) nor the irreducible representations is important. Usually, the totally symmetric representation $\boldsymbol{\Gamma}_1 \equiv \mathrm{A}_1$ appears as the first row of the table, while the identity operation, E, applies to the first column.

Table 12 The effect of each symmetry operation $\hat{R}$ of $\mathcal{C}_{2v}$ on the function $\phi = xy + zx + z$.

$\hat{R}$	$\hat{R}\phi$
E	$xy + zx + z$
$A = [2, 1] = C_2$	$(-x)(-y) + z(-x) + z$
$B = [2, 1]^{\#} = \sigma(xz)$	$x(-y) + zx + z$
$C = E^{\#} = \sigma(yz)$	$(-x)y + z(-x) + z$

totally symmetric (A_1). More generally, the various functions of first and second degrees in the coordinates are shown in the last column of Table 11 (see problem 16).

The example presented above is particularly simple because the irreducible representations are all of first order. However, in groups that contain noncommuting operations, degeneracies arise. In other words, some of the irreducible representations will be of second order or higher, depending on the symmetry of the system. This situation poses a problem in the application of Eq. (45), as it requires that all of the diagonal elements of each irreducible representation $\mathbf{\Gamma}^{(\gamma')}$ be known. In practice only the characters, the sum of the diagonal elements, are available from the character tables for the group in question.

If Eq. (45) is summed over the diagonal elements, a simplified projection operator can be defined, *viz.*

$$\begin{aligned}\hat{\mathcal{P}}^{(\gamma')} &= \sum_{t'} \hat{\mathcal{P}}^{(\gamma')}_{t',t'} = \frac{d_{\gamma'}}{g} \sum_{t'} \sum_{R} \Gamma^{(\gamma')\star}_{t',t'}(R)\hat{R} \\ &= \frac{d_{\gamma'}}{g} \sum_{R} \left[\sum_{t'} \Gamma^{(\gamma')\star}_{t',t'}(R) \right] \hat{R} = \frac{d_{\gamma'}}{g} \sum_{R} \chi^{(\gamma')}_{R} \hat{R}, \end{aligned} \tag{47}$$

where the order of summation has been changed and the quantity in brackets has been identified as the character.

An example of the application of Eq. (47) is provided by the group $\mathcal{C}_{3v}$, whose symmetry operations are defined by Eqs. (18). If the same arbitrary function, $\phi = xy + zx + z$, is used, the effect of each symmetry operation can be worked out, as shown in the last column of Table 13. With the use of the projection operator defined by Eq. (47) and the character table (Table 6), it is found (problem 16) that the coordinate z is totally symmetric (representation A_1). However, it is the sum $xy + zx$ that is preserved in the doubly degenerate representation, E. It should not be surprising that the functions xy and zx are projected as the sum, because it was the sum of the diagonal elements (the trace) of the irreducible representation that was employed in each case in the

Table 13 The effect of each symmetry operation of $\mathscr{C}_{3\mathrm{v}}$ on the function $\phi = xy + zx + z$.

Operation	x	y	z	$\hat{R}\phi$
E	x	y	z	$xy + zx + z$
$A = C_3^1$	$\frac{1}{2}(-y - \sqrt{3}x)$	$\frac{1}{2}(-y - \sqrt{3}x)$	z	$\frac{1}{2}(-x + \sqrt{3}y)\frac{1}{2}(-y - \sqrt{3}x) + \frac{1}{2}(-y - \sqrt{3}x)z + z$
$B = C_3^2$	$\frac{1}{2}(-x - \sqrt{3}y)$	$\frac{1}{2}(-y + \sqrt{3}x)$	z	$\frac{1}{2}(-x - \sqrt{3}y)\frac{1}{2}(-y + \sqrt{3}x) + \frac{1}{2}(-x - \sqrt{3}y)z + z$
$C = \sigma_1(yz)$	$-x$	y	z	$-xy - zx + z$
$D = \sigma_2$	$\frac{1}{2}(x - \sqrt{3}y)$	$\frac{1}{2}(-y - \sqrt{3}x)$		$\frac{1}{2}(x - \sqrt{3}y)\frac{1}{2}(-y - \sqrt{3}x) + \frac{1}{2}(x - \sqrt{3}y)z + z$
$F = \sigma_3$	$\frac{1}{2}(x + \sqrt{3}y)$	$\frac{1}{2}(-y + \sqrt{3}x)$		$\frac{1}{2}(x + \sqrt{3}y)\frac{1}{2}(-y + \sqrt{3}x) + \frac{1}{2}(x + \sqrt{3}y)z + z$

calculation. Thus, it is evident that the use of the simplified projection operator given by Eq. (47) provides less information than the general form [Eq. (45)]. However, in most practical applications it is quite sufficient.

8.12 HYBRIDIZATION OF ATOMIC ORBITALS

Of particular importance in structural chemistry is the concept of hybridization, that is, the construction of linear combinations of atomic orbitals that transform according to the symmetry of the structure. For the present, a simple illustration is provided by the hybridization of atomic orbitals in a molecule or complex ion of trigonal structure.

A planar molecule of point group $\mathscr{D}_{3\mathrm{h}}$ is shown in Fig. 5. The sigma orbitals σ_1, σ_2 and σ_3 represented there will be taken as the basis set. Application of the method developed in Section 8.9 yields the characters of the reducible representation given in Table 14. With the use of the magic formula [Eq. (37)] the structure of the reduced representation is of the form $\mathbf{\Gamma}_\sigma = A_1' \oplus E'$.

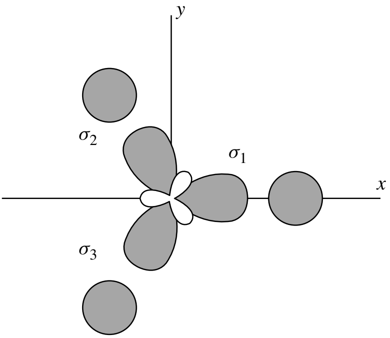


Fig. 5 Trigonal hybridization, sp^2.

Table 14 The characters of $\boldsymbol{\Gamma}_\sigma$ in $\mathscr{D}_{3\mathrm{h}}$.

$\mathscr{D}_{3\mathrm{h}}$	E	$2C_3$	$3C_2$	σ_h	$2S_3$	$3\sigma_\mathrm{v}$
$\chi_j(\sigma)$	3	0	1	3	0	1

Hence, the three appropriate linear combinations of the sigma orbitals can be constructed with the use of the projection operator defined by Eq. (47).

If σ_1 is chosen as the generating function, it yields the other two members of the set (as well as itself) under the symmetry operations of the point group. The function σ_1 is obviously the result of the identity operation, while C_3^1 and C_3^2 produce σ_2 and σ_3, respectively. These three symmetry operations are in fact sufficient to resolve the problem, although the reader can verify that if all of the operations of the group are employed, the same expression will be obtained (problem 17).

For the irreducible representation A_1' the symmetrized combination is easily found to be $\sigma_1 + \sigma_2 + \sigma_3$. The application of Eq. (47) for the representation E' yields $2\sigma_1 - \sigma_2 - \sigma_3$. As the simplified projection operator has been employed in this example, the second combination of species E' is not given directly. However, it is sufficient in this case to construct a second linear combination that is compatible with the symmetry E' and orthogonal to the first. A direct method to find the appropriate combination is to permute cyclically the functions obtained above, *viz.*

$$2\sigma_1 - \sigma_2 - \sigma_3 \rightarrow 2\sigma_3 - \sigma_1 - \sigma_2. \tag{48}$$

Clearly, the sum of the left-hand side of Eq. (48) and twice the right-hand side yields the combination $3(\sigma_3 - \sigma_2)$. Thus, the generating function has been eliminated and the resulting combination is orthogonal to the first one. The numerical factor involved is unimportant, as all of the resulting combinations will subsequently be correctly normalized. The result of this analysis is best expressed in matrix form, namely,

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = \begin{pmatrix} \dfrac{1}{\sqrt{3}} & \dfrac{1}{\sqrt{3}} & \dfrac{1}{\sqrt{3}} \\ \dfrac{2}{\sqrt{6}} & -\dfrac{1}{\sqrt{6}} & -\dfrac{1}{\sqrt{6}} \\ 0 & \dfrac{1}{2} & -\dfrac{1}{2} \end{pmatrix} \begin{pmatrix} s \\ p_x \\ p_y \end{pmatrix}. \tag{49}$$

where the appropriate normalization factors have been included. The 3×3 matrix that transforms the atomic orbitals into the symmetrized hybrids is then orthogonal.

It is usually of interest to express the atomic orbitals as functions of the hybrid orbitals. As the transformation matrix is orthogonal, its inverse is simply its transform. Thus, Eq. (49) becomes

$$\begin{pmatrix} s \\ p_x \\ p_y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{2} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix}, \tag{50}$$

which yields the expressions for the atomic orbitals as linear combinations of the hybrid orbitals for this particular structure.

The projection-operator technique will be employed in several examples presented in the following chapter and Chapter 12. For the quantitative interpretation of molecular spectra, both electronic and vibrational, molecular symmetry plays an all-important role. The correct linear combinations of electronic wavefunctions, as well as vibrational coordinates, are formed with the aid of the projection-operator method.

8.13 CRYSTAL SYMMETRY

The symmetry operations that characterize a point group were described in Section 8.5. The same operations can be used to specify the symmetry of a unit cell in a crystal. A given operation was represented by a matrix $\boldsymbol{R}$ which transforms a vector (a point in space) into another vector. A system of coordinates can always be chosen so that $\boldsymbol{R}$ is given by Eq. (7-104). However, in a crystal lattice the operations are restricted to those in which the final point is equivalent under simple translations to the initial one. Thus it is said that the unit cell of a crystal must be capable of “filling all space” *via* simple translations of its coordinates in the three (not necessarily Cartesian) directions. It can be shown that the corresponding mathematical condition is that the trace of $\boldsymbol{R}$ be an integer. Thus [Eq. (7-104)], if

$$\boldsymbol{R}\xi = \begin{bmatrix} cos\,\varphi & sin\,\varphi & 0 \\ -sin\,\varphi & cos\,\varphi & 0 \\ 0 & 0 & \pm 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}, \tag{51}$$

$Tr\boldsymbol{R} = 2\,cos\,\varphi \pm 1 = n$, a positive or negative integer, or zero. Then, $cos\,\varphi$ can only be integer or half-integer, and only axes of orders 1, 2, 3, 4 and 6 are possible (see problem 18). With this limitation it is found that only 32 groups can be formed from the operations that describe the symmetry of a unit cell. These point groups constitute the 32 crystal classes shown in Table 15.

Table 15 The crystallographic point groups (crystal classes).

Schönflies symbol	International symbol (Hermann-Maugin)	Crystal system
$\mathcal{C}_1$	1	Triclinic
$\mathcal{C}_\mathrm{i}$	$\bar{1}$	
$\mathcal{C}_\mathrm{s}$	m	Monoclinic
$\mathcal{C}_2$	2	
$\mathcal{C}_\mathrm{2h}$	$2/m$	
$\mathcal{C}_\mathrm{2v}$	mm	Orthorhombic
$\mathcal{D}_2$	222	
$\mathcal{D}_\mathrm{2h}$	mmm	
$\mathcal{C}_4$	4	Tetragonal
$\mathcal{S}_4$	$\bar{4}$	
$\mathcal{C}_\mathrm{4h}$	$4/m$	
$\mathcal{C}_\mathrm{4v}$	$4mm$	
$\mathcal{D}_\mathrm{2d}$	$\bar{4}2m$	
$\mathcal{D}_4$	422	
$\mathcal{D}_\mathrm{4h}$	$4/mmm$	
$\mathcal{C}_3$	3	Rhombohedral
$\mathcal{S}_6$	$\bar{3}$	
$\mathcal{C}_\mathrm{3v}$	$3m$	
$\mathcal{D}_3$	32	
$\mathcal{D}_\mathrm{3d}$	$\bar{3}m$	
$\mathcal{C}_\mathrm{3h}$	$\bar{6}$	Hexagonal
$\mathcal{C}_6$	6	
$\mathcal{C}_\mathrm{6h}$	$6/m$	
$\mathcal{D}_\mathrm{3h}$	$\bar{6}m2$	
$\mathcal{C}_\mathrm{6v}$	$6mm$	
$\mathcal{D}_6$	622	
$\mathcal{D}_\mathrm{6h}$	$6/mmm$	
$\mathcal{T}$	23	Cubic
$\mathcal{T}_\mathrm{h}$	$m3$	
$\mathcal{T}_\mathrm{d}$	$\bar{4}3m$	
$\mathcal{O}$	432	
$\mathcal{O}_\mathrm{h}$	$m3m$	

For the application to a crystal lattice Eq. (51) can be generalized to include the possibility of translation of a point in space by writing

$$\xi' = \boldsymbol{R}\xi + \boldsymbol{\tau}, \tag{52}$$

where the vector $\boldsymbol{\tau}$ represents a displacement or translation of the point ξ. Equation (52) is often written in the form

$$\xi' = \{\boldsymbol{R}|\boldsymbol{\tau}\}\xi. \tag{53}$$

Some general properties of operators of the type $\{\boldsymbol{R}|\boldsymbol{\tau}\}$ can be easily developed.

Consider the transformation of Eq. (52), followed by

$$\xi'' = \boldsymbol{R}'\xi' + \boldsymbol{\tau}' = \{\boldsymbol{R}'|\boldsymbol{\tau}'\}\xi'. \tag{54}$$

The result of the successive operations is then given by

$$\xi'' = \boldsymbol{R}'(\boldsymbol{R}\xi + \boldsymbol{\tau}) + \boldsymbol{\tau}' = \boldsymbol{R}'\boldsymbol{R}\xi + (\boldsymbol{R}'\boldsymbol{\tau} + \boldsymbol{\tau}'), \tag{55}$$

where the operation $\boldsymbol{R}'\boldsymbol{\tau} + \boldsymbol{\tau}'$ is also a translation. Thus, the product of two operations is defined by

$$\{\boldsymbol{R}'|\boldsymbol{\tau}'\}\{\boldsymbol{R}|\boldsymbol{\tau}\} = \{\boldsymbol{R}'\boldsymbol{R}|\boldsymbol{R}'\boldsymbol{\tau} + \boldsymbol{\tau}'\}. \tag{56}$$

Furthermore, the identity operation is $\{\boldsymbol{E}|0\}$. The inverse of the operator $\{\boldsymbol{R}|\boldsymbol{\tau}\}$ can be found from Eq. (56) by imposing the conditions $\boldsymbol{R}'\boldsymbol{R} = \boldsymbol{E}$ and $\boldsymbol{R}'|\boldsymbol{\tau} + \boldsymbol{\tau}' = \boldsymbol{0}$. Then,

$$\{\boldsymbol{R}|\boldsymbol{\tau}\}^{-1} = \{\boldsymbol{R}^{-1}| -\boldsymbol{R}^{-1}\boldsymbol{\tau}\}. \tag{57}$$

From the above properties it is evident that the set of operations $\{\boldsymbol{R}|\boldsymbol{\tau}\}$ forms a group, $\mathcal{S}$, the space group of the crystal. If the translation operations are the primitive translations $\{\boldsymbol{E}|\boldsymbol{\tau}_n\}$,

$$\boldsymbol{\tau}_n = n_1\boldsymbol{t}_1 + n_2\boldsymbol{t}_2 + n_3\boldsymbol{t}_3 \tag{58}$$

where n_1, n_2 and n_3 are integers. The vectors $\boldsymbol{t}_1, \boldsymbol{t}_2$ and $\boldsymbol{t}_3$ are the three linearly independent primitive lattice vectors. The points generated by the vectors constitute the lattice. It can be shown that only 14 different lattices, the Bravais lattices,* can be developed.

The set of translations $\{\boldsymbol{E}|\boldsymbol{\tau}_n\}$ forms a group $\mathcal{T}$, known as the translation group, an Abelian subgroup of $\mathcal{S}$. The space group can thus be written as

$$\mathcal{S} = \mathcal{U} \otimes \mathcal{T}, \tag{59}$$

---

*Auguste Bravais, French physicist (1811–1863).

where $\mathscr{U}$, the unit cell group, is isomorphic with the point group $\{\boldsymbol{R}|0\}$ that defines the crystal class.* This result provides the basis for the factor-group method, which is employed in the interpretation of the vibrational spectra of crystalline solids.

PROBLEMS

1. Develop Table 3, the multiplication table for the matrices given by Eqs. (2).

2. Verify Eqs. (5) and (6).

3. Verify Eqs. (11).

4. Show that the operations C, D and F (Table 3) are in the same class.

5. Identify the symmetry groups of the molecules represented in Fig. 1.

Ans. $\mathscr{C}_{2v}, \mathscr{D}_{3h}, \mathscr{D}_{6h}$

6. What are the symmetry groups of the possible conformations of the ethane molecule?

Ans. $\mathscr{D}_{3d}, \mathscr{D}_{3h}, \mathscr{D}_{3}$

7. The molecule LaB_6 occupies a cubic site in the crystal. Find the 48 symmetry operations and thus the point group of the molecule.

Ans. $\mathscr{O}_{h}$

8. The boron atom in natural abundance is approximately 80% ^{11}B and 20% ^{10}B. Identify all of the possible symmetry groups for LaB_6 with these isotopic species, *e.g.* $La^{11}B_6$, $La^{11}B_5{}^{10}B$, $La^{11}B_4{}^{10}B_2$, $La^{11}B_3{}^{10}B_3$ *etc.* A model is especially useful in this case.

Ans. $\mathscr{O}_{h}, \mathscr{C}_{4v}, \mathscr{C}_{2v}$ and $\mathscr{D}_{4h}, \mathscr{C}_{3v}$ and $\mathscr{C}_{2v}$ *etc.*

9. Show that the matrices given by Eqs. (14-17) represent the four-group (Table 2).

10. Demonstrate that the permutation operations defined by Eq. (18) form a group that is isomorphic with the group $\mathscr{C}_{3v}$.

11. Verify that the matrix $\boldsymbol{A}$ of Eq. (20) represents the operation $\begin{bmatrix} 3 & 1 & 2 \end{bmatrix} = C_3^1$ and find the corresponding representation of the operation $\begin{bmatrix} 2 & 3 & 1 \end{bmatrix}$.

12. Develop the analogous matrix representation of the operation $\begin{bmatrix} 2 & 1 & 3 \end{bmatrix}^{\#} = \sigma_3$.

---

*The direct product appearing in Eq. (59) is correct for those 73 space groups in which $\mathscr{U}$ is a point group (the symmorphic space groups). In the other 157 space groups, in which the crystal symmetry includes screw axes and/or glide planes, $\mathscr{U}$ is not a point group, and it is strictly speaking the semi-direct product that should be used in Eq. (59).

13. Derive the character table for the group $\mathscr{C}_{3\mathrm{v}}$ (Table 6).

14. With the use of the orthonormality of the characters, derive the magic formula [Eq. (37)].

15. Verify the values of $n^{(\gamma)}$ given in Table 9.

16. Verify the irreducible representations of the functions given in the last column of Table 11.

17. With the use of the projection operator derive the appropriate linear combinations sp^3 of the atomic orbitals under $\mathscr{D}_{3\mathrm{h}}$ symmetry.

18. Prove that if x', y', z' are integers, $cos\,\varphi$ in Eq. (51) can only be integer or half-integer.

9 Molecular Mechanics

A molecule is composed of a certain number N of nuclei and usually a much larger number of electrons. As the masses of the electrons and the nuclei are significantly different, the much lighter electrons move rapidly to create the so-called electron cloud which “sticks” the nuclei into relatively fixed equilibrium positions. The resulting geometry of the nuclear configuration is usually referred to as the molecular structure. The vibrational and rotational spectra of a molecule, as observed in its infrared absorption or emission and the Raman effect, are determined by this molecular geometry.

Within the framework of the Born–Oppenheimer approximation (see Chapter 12) the energy of a molecule can be written in the form

$$\varepsilon = \varepsilon_{elec} + \varepsilon_{nucl}, \tag{1}$$

where $\varepsilon_{\mathrm{elec}}$ is the energy associated with the electronic configuration of the molecule and $\varepsilon_{\mathrm{nucl}}$ is the energy of displacements of the nuclei. In general this approximation is an excellent one, hence, interaction terms can be neglected. The electronic energy serves as the effective potential function that governs the movement of the nuclei. As a first approximation, it is possible to decompose the energy associated with the nuclear displacements by writing

$$\varepsilon_{nucl} = \varepsilon_{trans} + \varepsilon_{rot} + \varepsilon_{vib} + \text{ interactions.} \tag{2}$$

Furthermore, it is convenient to consider separately the kinetic- and potential-energy contributions.

9.1 KINETIC ENERGY

The kinetic energy of a polyatomic molecule is a function of the atomic masses and the velocities of the atoms with respect to a space-fixed origin, O in Fig. 1. The center of mass of the molecule (cm) is located by the vector $\boldsymbol{R}$. The instantaneous position of each atom, α, with respect to the center of mass is specified by $\boldsymbol{r}_\alpha$ and the corresponding equilibrium position by $\boldsymbol{a}_\alpha$. Thus the vectors $\boldsymbol{\rho}_\alpha = \boldsymbol{r}_\alpha - \boldsymbol{a}_\alpha$ represent instantaneous displacements from equilibrium

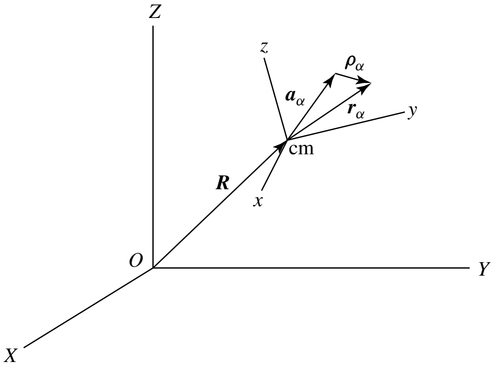


Fig. 1 Coordinate system used to specify atomic positions in space.

of the atoms $\alpha = 1, 2, \ldots, N$. The velocity of each atom with respect to the origin is given by

$$\boldsymbol{v}_\alpha = \dot{\boldsymbol{R}} + \dot{\boldsymbol{r}}_\alpha = \dot{\boldsymbol{R}} + \boldsymbol{v}_\alpha + \boldsymbol{\omega} \times \boldsymbol{r}_\alpha, \tag{3}$$

where $\boldsymbol{v}_\alpha$ is the velocity of atom α with respect to the center of mass and $\boldsymbol{\omega} \times \boldsymbol{r}_\alpha$ is the “velocity of following” that the atom must have to maintain its position in the molecule-fixed system x,y,z.

The total kinetic energy, T, resulting from the atomic displacements is given by

$$\begin{aligned}
2T &= \sum_\alpha m_\alpha \boldsymbol{v}_\alpha^2 \\
&= \dot{\boldsymbol{R}}^2 \sum_\alpha m_\alpha + \sum_\alpha m_\alpha (\boldsymbol{\omega} \times \boldsymbol{r}_\alpha) \cdot (\boldsymbol{\omega} \times \boldsymbol{r}_\alpha) + \sum_\alpha m_\alpha \boldsymbol{v}_\alpha^2 \\
&\quad + 2\dot{\boldsymbol{R}} \cdot \boldsymbol{\omega} \times \sum_\alpha m_\alpha \boldsymbol{r}_\alpha + 2\dot{\boldsymbol{R}} \cdot \sum_\alpha m_\alpha \boldsymbol{v}_\alpha + 2\boldsymbol{\omega} \cdot \sum_\alpha (m_\alpha \boldsymbol{r}_\alpha \times \boldsymbol{v}_\alpha),
\end{aligned} \tag{4}$$

where $\boldsymbol{\omega}$ is the angular velocity of the molecule-fixed coordinate system and m_α is the mass of the α^{th} atom.

For a free molecule, as in an ideal gas, Eq. (4) can be simplified with the use of the conditions

$$\sum_\alpha m_\alpha \boldsymbol{r}_\alpha = \sum_\alpha m_\alpha \boldsymbol{v}_\alpha = 0 \tag{5}$$

and

$$\sum_\alpha m_\alpha \boldsymbol{a}_\alpha \times \boldsymbol{v}_\alpha = 0. \tag{6}$$

Equation (5) expresses the conservation of linear momentum that defines the position of the center of mass of the molecule, while Eq. (6) is an approximate statement of the conservation of angular momentum of the system. These conditions, which are usually attributed to Eckart,* lead to the relation

$$2T = \dot{\boldsymbol{R}}^2 \sum_\alpha m_\alpha + \sum_\alpha m_\alpha (\boldsymbol{\omega} \times \boldsymbol{r}_\alpha) \cdot (\boldsymbol{\omega} \times \boldsymbol{r}_\alpha) + \sum_\alpha m_\alpha \boldsymbol{v}_\alpha^2 + 2\boldsymbol{\omega} \cdot \sum_\alpha m_\alpha (\boldsymbol{\rho}_\alpha \times \boldsymbol{v}_\alpha). \tag{7}$$

The first term on the right-hand side of Eq. (7) represents the kinetic energy of translation of the center of mass of the molecule. It will be shown in Chapter 10 that this term makes an important contribution to the partition function, although it is not significant in the vibrational spectroscopy of gases. The second term expresses the energy of overall rotation about the center of mass, cm, while the third term represents the total vibrational energy of the molecule resulting from the relative atomic displacements. The final term in Eq. (7) takes into account the so-called Coriolis interaction, the principal part of the kinetic coupling between the vibrational and rotational degrees of freedom.†

9.2 MOLECULAR ROTATION

For a free molecule the rotational energy is entirely kinetic, as given by the second term of Eq. (7). Furthermore, the problem can be simplified by assuming that the displacements $\boldsymbol{\rho}_\alpha$ are small. In this case the energy can be expressed as

$$\varepsilon_{rot} = \tfrac{1}{2}\tilde{\boldsymbol{\omega}} \boldsymbol{I} \boldsymbol{\omega}, \tag{8}$$

where $\boldsymbol{I}$ is the moment-of-inertia tensor. In the limit of infinitesimal vibrational amplitudes, the moment of inertia tensor becomes equal to that of a rigid rotator, *viz.*

$$\boldsymbol{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix}. \tag{9}$$

The Cartesian coordinates employed in Eq. (9) have been assumed to be attached to the rigid, rotating molecule. It should be noted that the moment-of-inertia tensor is symmetric with respect to the principal diagonal; thus,

---

*C. Eckart, American physicist (1904–). These conditions are sometimes attributed to Aaron Sayvetz, American physicist (1917–).

†Gaspard Coriolis, French engineer and mathematician (1792–1843).

$I_{xy} = I_{yx}$, *etc*. Equation (8) can then be written out in the form

$$\varepsilon_{rot} = \tfrac{1}{2}[I_{xx}\omega_x^2 + I_{yy}\omega_y^2 + I_{zz}\omega_z^2 + 2I_{xy}\omega_x\omega_y + 2I_{yz}\omega_y\omega_z + 2I_{zx}\omega_z\omega_x]. \quad (10)$$

It is in general possible to find a Cartesian coordinate system attached to the molecule such that the tensor $\boldsymbol{I}$ takes a diagonal form. In terms of these so-called principal axes Equation (10) is simplified in that all cross terms are eliminated. Namely,

$$\varepsilon_{rot} = \tfrac{1}{2}(I_a\omega_a^2 + I_b\omega_b^2 + I_c\omega_c^2). \quad (11)$$

The definition of the principal axes labeled a, b, c is usually dictated by the symmetry properties of the molecule (see Section 8.5). Then, the angular velocity can be represented by the vector

$$\boldsymbol{\omega} = \begin{pmatrix} \omega_a \\ \omega_b \\ \omega_c \end{pmatrix}.$$

The components of the angular momentum are defined by

$$\mathscr{L}_i \equiv \frac{\partial \varepsilon_{rot}}{\partial \omega_i}, \quad (12)$$

with $i = a, b, c$, which allows Eq. (11) to be written as

$$\varepsilon_{rot} = \frac{1}{2}\tilde{\boldsymbol{\omega}}\boldsymbol{I}\,\boldsymbol{\omega} = \frac{1}{2}\left(\frac{\mathscr{L}_a^2}{I_a} + \frac{\mathscr{L}_b^2}{I_b} + \frac{\mathscr{L}_c^2}{I_c}\right). \quad (13)$$

This expression for the classical rotational energy of a rigid body will now be developed in terms of Euler's angles.

9.2.1 Euler's angles

Consider the relations between a Cartesian coordinate system

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

attached to a molecule and another

$$\boldsymbol{X} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

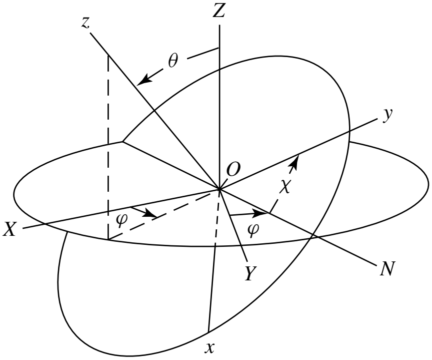


Fig. 2 Euler's angles.

fixed in space. These systems can be related by a matrix of direction cosines $\boldsymbol{R}$ such that $\boldsymbol{x} = \boldsymbol{R}\boldsymbol{X}$. It is convenient to express the direction cosines as functions of the Eulerian angles θ, φ, χ. They are defined in Fig. 2. The matrix $\boldsymbol{R}(\theta, \varphi, \chi)$ can be constructed by carrying out three successive rotations, as follows. The rotation about the vertical Z axis is described by a matrix

$$\boldsymbol{R}(\varphi) = \begin{pmatrix} cos\,\varphi & sin\,\varphi & 0 \\ -sin\,\varphi & cos\,\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{14}$$

a rotation about the Z axis by an angle φ [see Eq. (7-104)]. If it is followed by a rotation through an angle θ around the nodal line, ON, the corresponding matrix is of the form

$$\boldsymbol{R}(\theta) = \begin{pmatrix} cos\,\theta & 0 & -sin\,\theta \\ 0 & 1 & 0 \\ sin\,\theta & 0 & cos\,\theta \end{pmatrix}. \tag{15}$$

Finally, a rotation by an angle χ around the Oz axis is represented by the matrix

$$\boldsymbol{R}(\chi) = \begin{pmatrix} cos\,\chi & sin\,\chi & 0 \\ -sin\,\chi & cos\,\chi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{16}$$

These operations, if carried out in the correct order yield the desired transformation, *viz.*

$$\boldsymbol{R}(\theta, \varphi, \chi) = \boldsymbol{R}(\chi)\boldsymbol{R}(\theta)\boldsymbol{R}(\varphi). \tag{17}$$

It is not too difficult to show that the matrix product given by Eq. (17) leads to the elements shown in Table 1 (problem 5). These elements are of course

Table 1 The direction cosines expressed as functions of the Eulerian angles defined in Fig. 2.

	X	Y	Z
x	$\cos\theta\cos\varphi\cos\chi$ $-\sin\varphi\sin\chi$	$\cos\theta\sin\varphi\cos\chi$ $+\cos\varphi\sin\chi$	$-\sin\theta\cos\chi$
y	$-\cos\theta\cos\varphi\sin\chi$ $-\sin\varphi\cos\chi$	$-\cos\theta\sin\varphi\sin\chi$ $+\cos\varphi\cos\chi$	$\sin\theta\sin\chi$
z	$\sin\theta\cos\varphi$	$\sin\theta\sin\varphi$	$\cos\theta$

the direction cosines, which express the relation between the space-fixed and rotating coordinate system. The reader is warned that the results presented here depend on certain conventions, namely, right-handed coordinate systems and care in defining the direction of rotation of the Eulerian angles. Unfortunately, many variations on this theme are to be found in the scientific literature.

In the present application it is of interest to express the angular velocity of a rotating molecule in terms of the time derivatives of Euler's angles, $\dot{\varphi}, \dot{\theta}$ and $\dot{\chi}$. These quantities are directed along OZ, ON and Oz, respectively. Thus, the components of the angular velocity in the molecule-fixed system are given by

$$\begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} -\sin\theta\cos\chi & \sin\chi & 0 \\ \sin\theta\sin\chi & \cos\chi & 0 \\ \cos\theta & 0 & 1 \end{pmatrix} \begin{pmatrix} \dot{\varphi} \\ \dot{\theta} \\ \dot{\chi} \end{pmatrix}. \tag{18}$$

Note that the first column in the transformation matrix is just the last column of Table 1, while the second column is the same as the second column of $\boldsymbol{R}(\chi)$ [Eq. (16)]. Of course, as $\dot{\chi}$ is along the z axis, its coefficient is equal to one. Substitution of the angular velocity components given by Eq. (18) allows the rotational energy [Eq. (10)] to be expressed in terms of the velocities with respect to Euler's angles (see problem 7).

9.2.2 Classification of rotators

With the use of principal axes it becomes possible to classify the types of rotator as follows:

(i) Linear: $I_a = I_b \neq I_c = 0$
(ii) Spherical: $I_a = I_b = I_c \neq 0$
(iii) Symmetric: $I_a = I_b \neq I_c \neq 0$
(iv) Asymmetric: $I_a \neq I_b \neq I_c$.

The classical rotational energy of rigid body is given in the general case by Eq. (13). This expression is thus applicable to (iv), the asymmetric rotator. The other types of rotator represent special cases.

For the linear rotator of moment of inertia $I = I_a = I_b$ Eq. (13) becomes

$$\varepsilon_{rot} = \frac{\mathscr{L}^2}{2I}, \tag{19}$$

as $\mathscr{L}^2 = \mathscr{L}_a^2 + \mathscr{L}_b^2$ in this case. Similarly, for the spherical rotator with $I = I_a = I_b = I_c$ the energy is also given by Eq. (19). The energy expression for the symmetric rotator is somewhat more complicated, namely,

$$\varepsilon_{rot} = \frac{\mathscr{L}_a^2 + \mathscr{L}_b^2}{2I_a} + \frac{\mathscr{L}_c^2}{2I_c} = \frac{\mathscr{L}^2}{2I_a} + \frac{\mathscr{L}_c^2}{2}\left(\frac{1}{I_c} - \frac{1}{I_a}\right) \tag{20}$$

(see problem 8). Eq. (20) allows two types of symmetric rotator to be defined. If $I_c < I_a$, the rotator is elongated in the direction of the c axis. It is said to be a prolate top, as, say, an American football. The second term on the right-hand side of Eq. (20) is then positive. On the other hand, if $I_c > I_a$, the rotator is flattened with respect to the c direction, as the earth – or in the limit – a pancake. In this case the second term on the right-hand side of Eq. (20) is negative and the rotator is an oblate top. These two types of rotator can be distinguished in their rotational spectra by the contribution of the second term.

9.2.3 Angular momenta

It is often useful (and in quantum mechanics necessary) to express the rotational energy as a function of the components of the angular momentum. In the general case this transformation can be carried out as follows. The inverse of Eq. (18) is given by

$$\begin{pmatrix} \dot{\varphi} \\ \dot{\theta} \\ \dot{\chi} \end{pmatrix} = \begin{pmatrix} -\csc\theta\cos\chi & \csc\theta\sin\chi & 0 \\ \sin\chi & \cos\chi & 0 \\ \cot\theta\cos\chi & -\cot\theta\sin\chi & 1 \end{pmatrix} \begin{pmatrix} \omega_a \\ \omega_b \\ \omega_c \end{pmatrix}, \tag{21}$$

where the inverse matrix has been calculated with the use of Eq. (7-42) (see problem 9). The angular momenta with respect to the rotating axes can then be evaluated, as the chain rule yields

$$\begin{aligned} \mathscr{L}_a &= \frac{\partial\varepsilon_{rot}}{\partial\omega_a} = \frac{\partial\dot{\varphi}}{\partial\omega_a}\frac{\partial\varepsilon_{rot}}{\partial\dot{\varphi}} + \frac{\partial\dot{\theta}}{\partial\omega_a}\frac{\partial\varepsilon_{rot}}{\partial\dot{\theta}} + \frac{\partial\dot{\chi}}{\partial\omega_a}\frac{\partial\varepsilon_{rot}}{\partial\dot{\chi}} \\ &= \frac{\partial\dot{\varphi}}{\partial\omega_a}p_\varphi + \frac{\partial\dot{\theta}}{\partial\omega_b}p_\theta + \frac{\partial\dot{\chi}}{\partial\omega_c}p_\chi \end{aligned} \tag{22}$$

and similar relations for the other components of $\mathscr{L}$. The coefficients appearing in Eq. (22) are easily evaluated from Eq. (21). The angular momentum expressed in the rotating coordinates are then of the form

$$\mathscr{L} = \begin{pmatrix} \mathscr{L}_a \\ \mathscr{L}_b \\ \mathscr{L}_c \end{pmatrix} = \begin{pmatrix} -\csc\theta\cos\chi & \sin\chi & \cot\theta\cos\chi \\ \csc\theta\sin\chi & \cos\chi & -\cot\theta\sin\chi \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_\varphi \\ p_\theta \\ p_x \end{pmatrix}. \tag{23}$$

9.2.4 The symmetric top in quantum mechanics

To obtain the quantum-mechanical description of a symmetric rotator, each momentum component is replaced by the corresponding operator. Thus, $p_\varphi \rightarrow \hat{p}_\varphi = (\hbar/i)(\partial/\partial\varphi)$ *etc.*, as explained in Section 7.2. Then, from Eq. (23) the resulting expressions for the angular momentum operators are given explicitly by

$$\hat{\mathscr{L}}_a = \frac{\hbar}{i}\left(-\csc\theta\cos\chi\frac{\partial}{\partial\varphi} + \sin\chi\frac{\partial}{\partial\theta} + \cot\theta\cos\chi\frac{\partial}{\partial\chi}\right), \tag{24}$$

$$\hat{\mathscr{L}}_b = \frac{\hbar}{i}\left(\csc\theta\sin\chi\frac{\partial}{\partial\varphi} + \cos\chi\frac{\partial}{\partial\theta} - \cot\theta\sin\chi\frac{\partial}{\partial\chi}\right), \tag{25}$$

and

$$\hat{\mathscr{L}}_c = \frac{\hbar}{i}\frac{\partial}{\partial\chi} \tag{26}$$

(problem 11).

With these results for the angular-momentum operators it is possible to obtain the Hamiltonian for the rotation of a symmetric top by direct substitution in Eq. (13). The reader is warned that care must be taken in this substitution, as the order of the derivatives is to be rigorously respected. However, given sufficient patience one can show that the classical energy becomes the Hamiltonian operator in the form (problem 12)

$$\begin{aligned}\hat{H} = -\frac{\hbar^2}{2I_a}\Bigg[&\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\varphi}\frac{\partial^2}{\partial\varphi^2} + \left(\cot^2\theta + \frac{I_a}{I_c}\right)\frac{\partial^2}{\partial\chi^2} \\ &-\frac{2\cot\theta}{\sin\theta}\frac{\partial^2}{\partial\chi\partial\varphi}\Bigg].\end{aligned} \tag{27}$$

The resulting Schrödinger equation is $\hat{H}\psi(\theta,\varphi,\chi) = \varepsilon\psi(\theta,\varphi,\chi)$, where ε represents the eigenvalues of the Hamiltonian, *i.e.* the energies of the stationary rotational states of the system.

As the angular coordinates χ and φ enter Eq. (27) only as derivatives, they are cyclic coordinates (see Section 6.4.2). Therefore, the three independent variables can be separated, leading to wavefunctions of the form

$$\psi(\theta,\varphi,\chi) = \Theta(\theta)e^{iM\varphi}e^{iK\chi}, \tag{28}$$

where the condition for single-valuedness of the wavefunctions requires that both M and K be integer. Substitution of this expression in the wave equation yields the ordinary differential equation in θ (see problem 13).

$$\frac{1}{sin\,\theta}\frac{\partial}{\partial\theta}\left(sin\,\theta\frac{\partial\Theta(\theta)}{\partial\theta}\right)-\left[\frac{M^2}{sin^2\theta}+\left(cot^2\theta+\frac{I_a}{I_c}\right)K^2-\frac{2\,cot\,\theta}{sin\,\theta}KM\right.$$
$$\left.-\frac{2I_a}{\hbar^2}\varepsilon\right]\Theta(\theta)=0. \tag{29}$$

The trigonometric functions can be eliminated by introducing the change of independent variable $x=\frac{1}{2}(1-cos\,\theta)$. The result is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[x(1-x)\frac{\mathrm{d}u(x)}{\mathrm{d}x}\right]+\left\{C-\frac{[M+K(2x-1)]^2}{4x(1-x)}\right\}u(x)=0, \tag{30}$$

where $u(x)=\Theta(\theta)$ and $C=I_a\left(\frac{2\varepsilon}{\hbar^2}-\frac{K^2}{I_c}\right)$.

Equation (30) has two singularities, which are regular points, at $x=0$ and $x=1$. As shown in Section 5.5.2, they can be eliminated by the substitutions (see problem 15)

$$\Theta(\theta)=u(x)=x^{\frac{1}{2}|K-M|}(1-x)^{\frac{1}{2}|K+M|}y(x), \tag{31}$$

where the exponents have been determined by the two appropriate indicial equations. The resulting differential equation is of the form

$$x(1-x)\frac{\mathrm{d}^2y(x)}{\mathrm{d}x^2}+(p-qx)\frac{\mathrm{d}y(x)}{\mathrm{d}x}+ry(x)=0. \tag{32}$$

Comparison with Eq. (5-153) shows that the constants can be identified by

$$p=-\gamma=|K-M|+1, \tag{33}$$

$$q=-(1+\alpha+\beta)=|K+M|+|K-M|+2 \tag{34}$$

and

$$r=C+K^2-(\tfrac{1}{2}|K+M|+\tfrac{1}{2}|K+M|)(\tfrac{1}{2}|K+M|+\tfrac{1}{2}|K+M|+1). \tag{35}$$

The function $y(x)$ can now be developed in a power series following the method presented in Section 5.2.1. The recursion formula for the coefficients is then of the form

$$a_{n+1}=\left[\frac{n(n-1)+qn-r}{(n+1)(n+p)}\right]a_n. \tag{36}$$

This series is not acceptable unless it terminates at a specific term, say, $n = j$. Then, the numerator of Eq. (36) is equal to zero and the energy of the system is given by

$$\varepsilon_{rot} = \frac{\hbar^2}{2I_a} J(J+1) + \frac{\hbar^2 K^2}{2}\left(\frac{1}{I_c} - \frac{1}{I_a}\right). \tag{37}$$

The quantum number $J = j + \frac{1}{2}|K + M| + \frac{1}{2}|K - M|$ is thus equal to, or greater than, the larger of the two numbers $|K|$ or $|M|$, and is given by $J = 0, 1, 2, \ldots,$ with $K = 0, \pm 1, \pm 2, \ldots, \pm J$ and $M = 0, \pm 1, \pm 2, \ldots, \pm J$, independently. The degeneracy for this class of rotator is rather complicated, *viz.* $2J + 1$ if $K = 0$ and $2(2J + 1)$ if $|K| > 0$.

Comparison of Eq. (37) with Eq. (20) allows the angular momentum operators to be identified by

$$\hat{\mathscr{L}}^2 = \hat{\mathscr{L}}_a^2 + \hat{\mathscr{L}}_b^2 + \hat{\mathscr{L}}_c^2 = \hbar^2 J(J+1) \tag{38}$$

and

$$\hat{\mathscr{L}}_c = \hbar K. \tag{39}$$

Then, for a linear molecule, as $\hat{\mathscr{L}}_c = \hbar K = 0$, the rigid-rotator energy is given simply by

$$\varepsilon_{rot} = \frac{\hbar^2}{2I} J(J+1), \tag{40}$$

with the rotational quantum number equal to $J = 0, 1, 2, \ldots$. For the spherical top, K is not, in general equal to zero. However, as $I_a = I_c$, Eq. (37) becomes identical to Eq. (40). It should be noted, however, that for the spherical top the energy levels are $(2J + 1)^2$-fold degenerate.

Finally, for the determination of selection rules for rotational spectroscopy it is necessary to find the wavefunctions for this problem. This subject will be left for further development as given in numerous texts on molecular spectroscopy.

9.3 VIBRATIONAL ENERGY*

The vibrational energy of an isolated molecule will now be written as the sum of the kinetic and potential energies, namely,

$$\varepsilon_{vib} = T + V. \tag{41}$$

---

*The method developed in this section is usually attributed to Edgar Bright Wilson, Jr., American chemical physicist (1908–).

9.3.1 Kinetic energy

A set of $3N$ Cartesian displacement coordinates in the form

$$\xi = \begin{pmatrix} \Delta x_1 \\ \Delta y_1 \\ \Delta z_1 \\ \text{- - - -} \\ \Delta x_2 \\ \Delta y_2 \\ \Delta z_2 \\ \text{- - - -} \\ \cdot \\ \cdot \\ \cdot \\ \text{- - - -} \\ \Delta x_N \\ \Delta y_N \\ \Delta z_N \end{pmatrix} \tag{42}$$

can be employed to express the kinetic energy. If a diagonal matrix of the atomic masses is defined by

$$\boldsymbol{M} = \begin{pmatrix} m_1 & & & & & & & & & \\ & m_1 & & & & & & & & \\ & & m_1 & & & & \mathbf{0} & & & \\ & & & m_2 & & & & & & \\ & & & & m_2 & & & & & \\ & & & & & m_2 & & & & \\ & & & & & & \cdot & & & \\ & & & \mathbf{0} & & & & m_N & & \\ & & & & & & & & m_N & \\ & & & & & & & & & m_N \end{pmatrix}, \tag{43}$$

the vibrational kinetic energy in terms of these matrices takes the compact form

$$T_{vib} = \tfrac{1}{2}\tilde{\dot{\xi}}\boldsymbol{M}\dot{\xi} = \tfrac{1}{2}\tilde{\boldsymbol{p}}\boldsymbol{M}^{-1}\boldsymbol{p}, \tag{44}$$

where the dot over the vector defined by Eq. (42) indicates that the time derivative has been taken and the tilde represents the transpose operation. The second part of Eq. (44) is obtained from the definition of the momenta conjugate to the Cartesian displacement coordinates.

9.3.2 Internal coordinates: The *G* matrix

It is usually convenient to work with a set of internal displacement coordinates, $\boldsymbol{S}$, as they have chemical significance. In the limit of small amplitudes of atomic displacements, the two sets of coordinates are linearly related. Thus,

$$\boldsymbol{S} = \boldsymbol{B}\xi, \tag{45}$$

where the matrix $\boldsymbol{B}$ is determined by the equilibrium geometry of the molecule. The corresponding velocities are similarly related, *viz.*

$$\dot{\boldsymbol{S}} = \boldsymbol{B}\dot{\xi}. \tag{46}$$

The components of the momentum conjugate to the Cartesian displacement coordinates can be found by application of the chain rule,

$$p_j \equiv \frac{\partial T}{\partial \dot{\xi}_j} = \sum_t \frac{\partial \dot{S}_t}{\partial \dot{\xi}_j} \cdot \frac{\partial T}{\partial \dot{S}_t}. \tag{47}$$

From Eq. (46) the elements of $\boldsymbol{B}$ are given by $B_{tj} = \partial \dot{S}_t / \partial \dot{\xi}_j$, which allows Eq. (47) to be written as

$$\boldsymbol{p} = \tilde{\boldsymbol{B}}\boldsymbol{P}. \tag{48}$$

Here the elements of $\boldsymbol{P}$ are the momenta conjugate to the internal coordinates, thus $P_t \equiv \partial T / \partial \dot{S}_t$ and $\dot{S}_t = \partial T / \partial P_t$. Substitution of Eq. (48) in Eq. (44) leads to the vibrational kinetic energy in the form

$$T_{vib} = \tfrac{1}{2}\tilde{\boldsymbol{P}}\boldsymbol{B}\boldsymbol{M}^{-1}\tilde{\boldsymbol{B}}\boldsymbol{P} = \tfrac{1}{2}\tilde{\boldsymbol{P}}\boldsymbol{G}\boldsymbol{P}, \tag{49}$$

where the matrix $\boldsymbol{G}$ is defined by $\boldsymbol{G} \equiv \boldsymbol{B}\boldsymbol{M}^{-1}\tilde{\boldsymbol{B}}$. Although, in principle, the $\boldsymbol{G}$ matrix can be constructed with a knowledge of the atomic masses and the molecular geometry, tables of its elements are available that were developed many years ago. In present-day computer programs it is constructed and employed directly to calculate the vibrational frequencies and the specific forms of the normal modes of vibration, as summarized in the following sections.

The vibrational kinetic energy can also be expressed in terms of the velocities in internal coordinates by taking the partial derivatives of Eq. (49). Thus, $\dot{\boldsymbol{S}} = \boldsymbol{G}\boldsymbol{P}$ and, as $\boldsymbol{G}$ is square and nonsingular, $\boldsymbol{P} = \boldsymbol{G}^{-1}\dot{\boldsymbol{S}}$ and its transpose becomes $\tilde{\boldsymbol{P}} = \tilde{\dot{\boldsymbol{S}}}\boldsymbol{G}^{-1}$ (It should be noted that both $\boldsymbol{G}$ and its inverse are symmetric). With these substitutions Eq. (49) becomes

$$T_{vib} = \tfrac{1}{2}\tilde{\dot{\boldsymbol{S}}}\boldsymbol{G}^{-1}\dot{\boldsymbol{S}}. \tag{50}$$

Because of the form of Eq. (50), $\boldsymbol{G}$ is often referred to as the “inverse kinetic-energy matrix”.

9.3.3 Potential energy

Although the equilibrium configuration of a molecule can usually be specified, at ordinary temperatures, all of the atoms undergo oscillatory motions. The forces between the atoms in the molecule are described by a Taylor series of the intramolecular potential function in the internal coordinates. This function can then be written in the form

$$V = \tfrac{1}{2}\sum_{t,t'}\left(\frac{\partial^2 V}{\partial S_t \partial S_{t'}}\right)_0 S_t S_{t'} = \tfrac{1}{2}\sum_{t,t'} f_{t,t'} S_t S_{t'}, \tag{51}$$

an expression that includes the first nonzero terms in the expansion of the potential function with respect to the equilibrium configuration. Clearly, Eq. (51) is a generalization of the harmonic potential for the diatomic molecule (see Fig. 6–6). Although higher terms in this series can be added to take into account the anharmonicity of the molecular vibrations, the harmonic approximation is employed in most practical calculations of the vibrational frequencies of a polyatomic molecule.

Equation (51) can be written in matrix form as

$$V = \tfrac{1}{2}\tilde{\boldsymbol{S}}\boldsymbol{F}\boldsymbol{S}, \tag{52}$$

where $\boldsymbol{F}$ is a symmetric matrix of the force constants. Although the force constants can in principle be calculated from the molecular electronic wavefunctions, the results are not yet quantitatively reliable. These quantities are usually treated as variable parameters in the vibrational problem. Their evaluation from experimental spectroscopic data is of importance, both as a means of testing the results of *ab initio* electronic calculations and for their chemical significance. For example the knowledge of the value of a force constant associated with the stretching of a chemical bond provides some insight into the nature of the bond, including semiquantitative measures of both its length and its dissociation energy.

9.3.4 Normal coordinates

It is now fundamental to define the normal coordinates of this vibrational system – that is to say, the nuclear displacements in a polyatomic molecule. Again in the limit of small amplitudes of vibration, the normal coordinates in the form of the vector $\boldsymbol{Q}$, are related to the internal coordinates by a linear transformation, *viz.*

$$\boldsymbol{S} = \boldsymbol{L}\boldsymbol{Q}. \tag{53}$$

The kinetic and potential energies given by Eqs. (50) and (52) become

$$T_{vib} = \tfrac{1}{2}\tilde{\dot{\boldsymbol{S}}}\boldsymbol{G}^{-1}\dot{\boldsymbol{S}} = \tfrac{1}{2}\tilde{\dot{\boldsymbol{Q}}}\tilde{\boldsymbol{L}}\boldsymbol{G}^{-1}\boldsymbol{L}\dot{\boldsymbol{Q}} = \tfrac{1}{2}\tilde{\dot{\boldsymbol{Q}}}\boldsymbol{E}\dot{\boldsymbol{Q}}, \tag{54}$$

and

$$V = \tfrac{1}{2}\tilde{\boldsymbol{S}}\boldsymbol{F}\boldsymbol{S} = \tfrac{1}{2}\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{L}}\boldsymbol{F}\boldsymbol{L}\boldsymbol{Q} = \tfrac{1}{2}\tilde{\boldsymbol{Q}}\boldsymbol{\Lambda}\boldsymbol{Q}, \tag{55}$$

respectively. The matrices $\boldsymbol{E}$ and $\boldsymbol{\Lambda}$ are both diagonal. The former is the unit matrix, while the latter is composed of elements $\lambda_k = 4\pi^2\nu_k^2$. The vibrational frequencies are given by ν_k, where the subscript k identifies the normal mode of vibration. An inspection of Eqs. (54) and (55) identifies the role of the normal coordinates, namely, to eliminate simultaneously all cross-terms in the expressions for the kinetic and potential energies. This condition can be taken as the definition of the normal modes of vibration.

9.3.5 Secular determinant

The last step in the calculation of the frequencies of molecular vibrations, as observed in the infrared spectra, is carried out by combining Eqs. (54) and (55). The vibrational energy of a polyatomic molecule is then given in this, the harmonic approximation, by

$$\varepsilon_{vib} = T_{vib} + V = \tfrac{1}{2}\tilde{\dot{\boldsymbol{Q}}}\boldsymbol{E}\dot{\boldsymbol{Q}} + \tfrac{1}{2}\tilde{\boldsymbol{Q}}\boldsymbol{\Lambda}\boldsymbol{Q}. \tag{56}$$

From Eqs. (54) and (55) it is apparent that $\tilde{\boldsymbol{L}}\boldsymbol{G}^{-1}\boldsymbol{L} = \boldsymbol{E}$ and $\tilde{\boldsymbol{L}}\boldsymbol{F}\boldsymbol{L} = \boldsymbol{\Lambda}$. With the definition of the $\boldsymbol{G}$ matrix, $\tilde{\boldsymbol{L}} = \boldsymbol{L}^{-1}\boldsymbol{G}$. Therefore,

$$\boldsymbol{L}^{-1}\boldsymbol{G}\boldsymbol{F}\boldsymbol{L} = \boldsymbol{\Lambda}. \tag{57}$$

Hence, as is often stated, the determination of the normal coordinates is equivalent to the successful search for a matrix $\boldsymbol{L}$ that diagonalizes the product $\boldsymbol{GF}$ *via* a similarity transformation. This system of linear, simultaneous homogeneous equations can be written in the form

$$\sum_{t'}[(\boldsymbol{GF})_{tt'} - \delta_{tt'}\lambda_k]L_{t'k} = 0, \tag{58}$$

where $\delta_{t,t'}$ is the Kronecker delta.

This set of equations for the elements of $\boldsymbol{L}$ can be resolved by application of Cramer's rule. Then, a nontrivial solution exists only if the determinant of the coefficients vanishes, or

$$|\boldsymbol{GF} - \boldsymbol{E}\lambda_k| = 0. \tag{59}$$

This condition on the so-called secular determinant is the basis of the vibrational problem. The roots of Eq. (59), λ_k, are the eigenvalues of the matrix product $\boldsymbol{GF}$, while the columns of $\boldsymbol{L}$, the eigenvectors, determine the forms of the normal modes of vibration. These relatively abstract relations become more evident with the consideration of an example.

9.3.6 An example: The water molecule

As a simple illustration of the development of the secular determinant, consider the water molecule. A reasonable set of internal coordinates consists of the changes in lengths of the two bonds and the variation in the bond angle. Thus, from Eq. (45) and Fig. 2,

$$\boldsymbol{S} = \begin{pmatrix} \Delta r_1 \\ \Delta r_2 \\ \Delta\alpha \end{pmatrix} = \boldsymbol{B}\xi$$

$$= \left(\begin{array}{ccc|ccc|ccc} 0 & -s & -c & 0 & 0 & 0 & 0 & s & c \\ 0 & 0 & 0 & 0 & s & -c & 0 & -s & c \\ 0 & -c/r & s/r & 0 & c/r & s/r & 0 & 0 & -2s/r \end{array}\right) \left(\begin{array}{c} \Delta x_1 \\ \Delta y_1 \\ \Delta z_1 \\ \hline \Delta x_2 \\ \Delta y_2 \\ \Delta z_2 \\ \hline \Delta x_3 \\ \Delta y_3 \\ \Delta z_3 \end{array}\right), \tag{60}$$

where $s = sin\,\alpha/2, c = cos\,\alpha/2$ (with α the equilibrium bond angle) and r is the equilibrium O–H bond length, as shown in Fig. 3. The masses of the hydrogen and oxygen atoms are arranged according to Eq. (42) in the order established by the vector ξ namely,

$$\boldsymbol{M} = \begin{pmatrix} m_{\mathrm{H}} & & & & & & & & \\ & m_{\mathrm{H}} & & & & & & & \\ & & m_{\mathrm{H}} & & & \mathbf{0} & & & \\ & & & m_{\mathrm{H}} & & & & & \\ & & & & m_{\mathrm{H}} & & & & \\ & & & & & m_{\mathrm{H}} & & & \\ & & & \mathbf{0} & & & m_{\mathrm{O}} & & \\ & & & & & & & m_{\mathrm{O}} & \\ & & & & & & & & m_{\mathrm{O}} \end{pmatrix}. \tag{61}$$

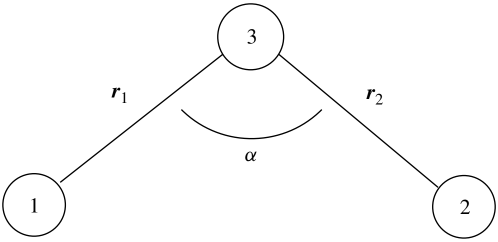


Fig. 3 Structure of the water molecule.

The $\boldsymbol{G}$ matrix, as calculated from its definition, becomes

$$\boldsymbol{G} = \boldsymbol{B}\boldsymbol{M}^{-1}\tilde{\boldsymbol{B}} = \begin{pmatrix} \mu_H + \mu_O & \mu_O \cos\alpha & -\dfrac{\mu_O \sin\alpha}{r} \\ \mu_O \cos\alpha & \mu_H + \mu_O & -\dfrac{\mu_O \sin\alpha}{r} \\ -\dfrac{\mu_O \sin\alpha}{r} & -\dfrac{\mu_O \sin\alpha}{r} & \dfrac{2\mu_H + 2\mu_O(1 - \cos\alpha)}{r^2} \end{pmatrix}, \tag{62}$$

where $\mu_H = 1/m_H$ and $\mu_O = 1/m_O$ (problem 19).

The force-constant matrix based on internal coordinates is of the form

$$\boldsymbol{F} = \begin{pmatrix} f_r & f_{rr} & f_{r\alpha} \\ f_{rr} & f_r & f_{r\alpha} \\ f_{r\alpha} & f_{r\alpha} & f_{\alpha\alpha} \end{pmatrix}, \tag{63}$$

where f_r is the coefficient of $(\Delta r_1)^2$ and $(\Delta r_2)^2$, and f_α is the coefficient of $(\Delta\alpha)^2$ in the potential energy expression. These constants are referred to as the principal or valence force constants. The off-diagonal constant f_{rr} represents an interaction constant which is the coefficient of $(\Delta r_1 \Delta r_2)$ and $(\Delta r_2 \Delta r_1)$ in the potential energy expression, while the interaction constant $f_{r\alpha}$ is the coefficient of all four terms of the type $(\Delta r_1 \Delta\alpha)$. By taking the product of the above matrices $\boldsymbol{G}$ and $\boldsymbol{F}$, the secular equation in the form of Eq. (36) can be found, albeit rather complicated (problem 20).

In the construction of the matrix $\boldsymbol{F}$ of Eq. (63), the symmetrical equivalence of the two O–H bonds was taken into account. Nevertheless, it contains four independent force constants. As the water molecule has but three fundamental vibrational frequencies, at least one interaction constant must be neglected or some other constraint introduced. If all of the off-diagonal elements of $\boldsymbol{F}$ are neglected, the two principal constants, f_r and f_α constitute the valence force field for this molecule. However, to reproduce the three observed vibrational frequencies this force field must be modified to include the interaction constant

f_{rr}. This example is continued in the following section, to illustrate the use of group theory to simplify the problem.

9.3.7 Symmetry coordinates

The characters χ_j for the examples in the previous section were calculated following the method described in Section 8.9, that is, on the basis of Cartesian displacement coordinates. Alternatively, it is often desirable to employ a set of internal coordinates as the basis. However, they must be well chosen so that they are sufficient to describe the vibrational degrees of freedom of the molecule and that they are linearly independent. The latter condition is necessary to avoid the problem of redundancy. Even when properly chosen, the internal coordinates still do not usually transform following the symmetry of the molecule. Once again, the water molecule provides a very simple example of this problem.

The internal coordinates for the water molecule are chosen as changes in the structural parameters defined in Fig. 3. The effect of each symmetry operation of the symmetry group $\mathcal{C}_{2\mathrm{v}}$ on these internal coordinates is specified in Table 2. Clearly, the internal coordinate $\Delta\alpha$ is totally symmetric, as the characters $\chi_j(\Delta\alpha)$ correspond to those given for the irreducible representation (IR) A_1. On the other hand, the characters $\chi_j(\Delta r)$, as shown, can not be identified with a specific IR. By inspection of Table 2, however, it is apparent that the direct sum $\mathrm{A}_1 \oplus \mathrm{B}_2$ corresponds to the correct symmetry of these coordinates. In more complicated cases the magic formula can always be employed to achieve the correct reduction of the representation in question.

From the above analysis, as shown in the last column of Table 2, it should be obvious that two linear combinations of the internal coordinates Δr can be formed whose symmetry corresponds to the IRs A_1 and B_2. They are $\Delta r_1 + \Delta r_2(\mathrm{A}_1)$ and $\Delta r_1 - \Delta r_2(\mathrm{B}_2)$, as can be verified by inspection of the character table. The choice of the correct linear combinations is in this case simple. However, more generally, it can be made by application of the projection-operator method described in Section 8.11.

Table 2 Symmetry of the internal coordinates for the water molecule.

$\mathcal{C}_{2\mathrm{v}}$	E	C_2	$\sigma(xz)$	$\sigma(yz)$	$n^{(\gamma)}(\Delta\alpha)$	$n^{(\gamma)}(\Delta r)$
A_1	1	1	1	1	1	1
A_2	1	1	−1	−1	0	0
B_1	1	−1	1	−1	0	0
B_2	1	−1	−1	1	0	1
$\chi_j(\Delta\alpha)$	1	1	1	1		
$\chi_j(\Delta r)$	2	0	0	2		

Coordinates such as these, which have the symmetry properties of the point group are known as symmetry coordinates. As they transform in the same manner as the IRs when used as basis coordinates, they factor the secular determinant into block-diagonal form. Thus, while normal coordinates must be found to diagonalize the secular determinant, the factorization resulting from the use of symmetry coordinates often provides considerable simplification of the vibrational problem. Furthermore, symmetry coordinates can be chosen *a priori* by a simple analysis of the molecular structure.

In the example considered above, $\Delta r_1 - \Delta r_2$ is the only symmetry coordinate of species B_2. Thus, it results in a factor of degree one in the completely reduced secular determinant. It is therefore a normal coordinate. On the other hand, the two normal coordinates of species A_1 are linear combinations of the symmetry coordinates $\Delta\alpha$ and $\Delta r_1 + \Delta r_2$. They can only be found by solution of the secular equations.

It is usually convenient to normalize the symmetry coordinates. Hence, for the example considered here, the three symmetry coordinates take the form

$$\text{Species } A_1 : \begin{cases} \mathsf{S}_1 = \Delta\alpha \\ \mathsf{S}_2 = \frac{1}{\sqrt{2}}(\Delta r_1 + \Delta r_2) \end{cases} \tag{64}$$

and

$$\text{Species } B_2 : \mathsf{S}_3 = \frac{1}{\sqrt{2}}(\Delta r_1 - \Delta r_2), \tag{65}$$

where the normalizing factors $\mathcal{N}_p$ are determined by the condition

$$\sum_t \mathcal{N}_t^2 = 1 \tag{66}$$

for each symmetry coordinate S_p. In general form, the relations given by Eqs. (64) and (65) are expressed by the matrix relation

$$\mathbf{\mathsf{S}} = \boldsymbol{U}\boldsymbol{S}\,, \tag{67}$$

which for the above example becomes

$$\begin{pmatrix} \mathsf{S}_1 \\ \mathsf{S}_2 \\ \mathsf{S}_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} \Delta r_1 \\ \Delta r_2 \\ \Delta\alpha \end{pmatrix}. \tag{68}$$

It should be noted that when the symmetry coordinates have been normalized, $\boldsymbol{U}$ is orthogonal; that is, $\tilde{\boldsymbol{U}} = \boldsymbol{U}^{-1}$.

9.3.8 Application to molecular vibrations

To see how use is made of symmetry coordinates as the bases of the vibrational problem, reconsider the kinetic and potential energies as given earlier, *e.g.*

$$2T = \tilde{\dot{\boldsymbol{S}}}\boldsymbol{G}^{-1}\boldsymbol{S} = \tilde{\dot{\boldsymbol{Q}}}\boldsymbol{E}\dot{\boldsymbol{Q}} \tag{69}$$

and

$$2V = \tilde{\boldsymbol{S}}\boldsymbol{F}\boldsymbol{S} = \tilde{\boldsymbol{Q}}\boldsymbol{\Lambda}\boldsymbol{Q}. \tag{70}$$

From Equation (67)

$$\boldsymbol{S} = \boldsymbol{U}^{-1}\mathbf{S} = \tilde{\boldsymbol{U}}\mathbf{S}; \tag{71}$$

hence,

$$2T = \tilde{\dot{\mathbf{S}}}\mathbf{G}^{-1}\dot{\mathbf{S}} = \tilde{\dot{\boldsymbol{Q}}}\boldsymbol{E}\dot{\boldsymbol{Q}} \tag{72}$$

and

$$2V = \tilde{\mathbf{S}}\mathbf{F}\mathbf{S} = \tilde{\boldsymbol{Q}}\boldsymbol{\Lambda}\boldsymbol{Q}. \tag{73}$$

In Eqs. (72) and (73) the definitions $\mathbf{G} = \boldsymbol{U}\boldsymbol{G}\tilde{\boldsymbol{U}}$ and $\mathbf{F} = \boldsymbol{U}\boldsymbol{F}\tilde{\boldsymbol{U}}$ have been introduced. These equations lead to the secular determinant as before,

$$|\mathbf{G}\mathbf{F} - \boldsymbol{E}\lambda_k| = 0. \tag{74}$$

The eigenvectors $\mathbf{L}$ with this basis can be found from the relation $\mathbf{L} = \boldsymbol{U}\boldsymbol{L}$.

Returning now to the example of the water molecule, with $\boldsymbol{F}$ given by Eq. (63) and $\boldsymbol{U}$ in Eq. (68), the transformation $\boldsymbol{U}\boldsymbol{F}\tilde{\boldsymbol{U}}$ gives

$$\mathbf{F} = \left(\begin{array}{cc:c} f_\alpha & \sqrt{2}f_{r\alpha} & 0 \\ \sqrt{2}f_{r\alpha} & f_r + f_{rr} & 0 \\ \hdashline 0 & 0 & f_r - f_{rr} \end{array}\right). \tag{75}$$

Similarly, $\boldsymbol{U}\boldsymbol{G}\tilde{\boldsymbol{U}}$ leads to

$$\mathbf{G} = \left(\begin{array}{cc:c} \dfrac{2\mu_{\mathrm{H}} + 2\mu_{\mathrm{O}}(1 - \cos\alpha)}{r^2} & -\dfrac{\sqrt{2}\mu_{\mathrm{O}}\sin\alpha}{r} & 0 \\ -\dfrac{\sqrt{2}\mu_{\mathrm{O}}\sin\alpha}{r} & \mu_{\mathrm{H}} + \mu_{\mathrm{O}}(1 + \cos\alpha) & 0 \\ \hdashline 0 & 0 & \mu_{\mathrm{H}} + \mu_{\mathrm{O}}(1 - \cos\alpha) \end{array}\right). \tag{76}$$

As both $\mathbf{F}$ and $\mathbf{G}$ are partitioned by the use of symmetry coordinates, the secular determinant is factored accordingly. The problem of calculating the vibrational frequencies is thus divided into two parts: solution of a linear equation for the single frequency of species B_2 and of a quadratic equation for the pair of frequencies of species A_1.

It is often of interest to calculate force constants from observed vibrational frequencies. However, it is not generally possible to derive analytical expressions for the force constants as functions of the frequencies and the molecular geometry. The calculation is necessarily an iterative one. Starting with a set of assumed force constants – usually obtained by analogy with similar bonds in other molecules – the values are refined until a suitable set is found. The set that yields the best agreement between calculated and observed frequencies constitutes the accepted force field for the molecule.

From the above example it is apparent that there may be (and usually are) more unknown force constants than observed frequencies. If no additional sources of data are available, it is necessary to make some assumptions to simplify the force field. Often all or some of the off-diagonal elements in the $\boldsymbol{F}$ matrix [Eq. (63)] are neglected, leading to the so-called valence force field (VFF) or modified valence force field (MVFF), respectively.

Assuming that a reasonable force field is known, the solution of the above equations to obtain the vibrational frequencies of water is not difficult. However, in more complicated molecules it becomes very rapidly a formidable one. If there are N atoms in the molecule, there are $3N$ total degrees of freedom and $3N-6$ for the vibrational frequencies. The molecular symmetry can often aid in simplifying the calculations, although in large molecules there may be no true symmetry. In some cases the notion of local symmetry can be introduced to simplify the calculation of vibrational frequencies and the corresponding forms of the normal modes of vibration.

Regardless of the force field chosen, the calculation of vibrational frequencies by the method outlined above is based on the harmonic approximation. Tabulated values of force constants can be used to calculate vibrational frequencies, for example, of molecules whose vibrational spectra have not been observed. However, as anharmonicities have been neglected in the above analysis, the resulting frequency values are often no better than $\pm 5\%$ with respect to those observed.

9.3.9 Form of normal modes

For the example of the water molecule it is of interest to calculate the forms of the vibrational modes, as obtained from the evaluation of the matrix $\mathbf{L} = \boldsymbol{UL}$. The results can be presented most simply as shown in Fig. 4. The calculation of the specific form of the normal modes is complicated, although with the aid of current computer programs it becomes routine – at least for relatively simple molecules.

It is apparent from Fig. 4 that the normal modes of vibration of the water molecule, as calculated from the eigenvectors, can be described approximately as a symmetrical stretching vibration (ν_1) and a symmetrical bending vibration

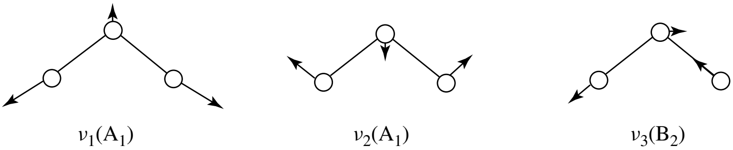


Fig. 4 The normal modes of vibration of the water molecule.

(ν_2). It is to be emphasized that the two normal modes of species A_1 are linear combinations of the two internal symmetry coordinates given by Eq. (64). On the other hand, as there is only one symmetry coordinate of species B_2 in this case, it is identical to the normal coordinate shown for ν_3 and can be correctly described as an antisymmetric stretching vibration.

It has been shown that the potential energy distribution provides an approximate method to evaluate the relative contribution of each symmetry coordinate to a given normal mode of vibration. From the definition of the symmetry coordinates, the relation

$$\tilde{\mathsf{L}}\mathsf{F}\mathsf{L} = \boldsymbol{\Lambda} \tag{77}$$

or

$$\sum_{pp'} \mathsf{L}_{pk}\mathsf{L}_{p'k}\mathsf{F}_{pp'} = \lambda_k \qquad k = 1, 2, \ldots, 3N - 6 \tag{78}$$

can be derived. As the diagonal elements of F are usually at least an order of magnitude greater than the off-diagonal elements, the relation

$$\frac{1}{\lambda_k}\sum_{p} \mathsf{L}_{pk}^2\mathsf{F}_{pp} \approx 1 \qquad k = 1, 2, \ldots, 3N - 6 \tag{79}$$

is approximately equivalent to Eq. (78). Each term on the left-hand side of Eq. (79) represents the approximate contribution of each symmetry coordinate S_p to the normal coordinate Q_k. The evaluation of this potential-energy distribution, as well as the form of the normal vibrations, is usually carried out by current computer programs for the calculation of vibrational frequencies.

To find the atomic displacements for each normal mode of vibration, use is made of Eqs. (45) and (53),

$$\boldsymbol{S} = \boldsymbol{B}\xi = \boldsymbol{L}\boldsymbol{Q}. \tag{80}$$

The matrix of Cartesian displacements ξ could in principle be found by multiplying Eq. (80) by the inverse of $\boldsymbol{B}$. However, $\boldsymbol{B}$ is not, in general square and nonsingular, and hence cannot be inverted. Nevertheless, as $\boldsymbol{G} \equiv \boldsymbol{B}\boldsymbol{M}^{-1}\tilde{\boldsymbol{B}}$, it is evident that the product $\boldsymbol{M}^{-1}\tilde{\boldsymbol{B}}\boldsymbol{G}^{-1}$ serves as the inverse of $\boldsymbol{B}$ in the sense that $\boldsymbol{B}(\boldsymbol{M}^{-1}\tilde{\boldsymbol{B}}\boldsymbol{G}^{-1}) = (\boldsymbol{M}^{-1}\tilde{\boldsymbol{B}}\boldsymbol{G}^{-1})\boldsymbol{B} = \boldsymbol{E}$. Hence, premultiplying Eq. (80) by $\boldsymbol{M}^{-1}\tilde{\boldsymbol{B}}\boldsymbol{G}^{-1}$ yields the result

$$\xi = \boldsymbol{M}^{-1}\tilde{\boldsymbol{B}}\boldsymbol{G}^{-1}\boldsymbol{S} = (\boldsymbol{M}^{-1}\tilde{\boldsymbol{B}}\boldsymbol{G}^{-1}\boldsymbol{L})\boldsymbol{Q}, \tag{81}$$

which is the desired transformation. Numerical solutions of the vibrational problem usually include calculations of the normal modes in Cartesian displacement coordinates based on Eq. (81).

9.4 NONRIGID MOLECULES

Many molecules have more than one well-defined structure – or even none. If there is more than one equilibrium structure the passage from one to another can take place because of the "tunnel effect", although it may be impossible from a purely classical point of view. The best known example is certainly the ammonia molecule, NH_3.

9.4.1 Molecular inversion*

The equilibrium structure of ammonia is that of a pyramid whose equilateral-triangular base is defined by the three protons. Its molecular symmetry is described by $\mathcal{C}_{3v}$, a group of order six (see Table 8-6). In its ground vibrational state at ordinary temperatures the molecule does not have enough kinetic energy to allow the protons to pass to equivalent positions on the other side of the nitrogen atom, as shown in Fig. 4. This inversion process, which is associated with the symmetrical (stretching-bending) vibrational mode (A_1), takes place by tunneling through the potential barrier (see Section 5.4.3).

The presence of the potential barrier in the ammonia molecule results in splitting of the vibrational energy levels, as shown in Fig. 5. The separation between the two components of the first level is equal to 23.87 GHz (≈ 0.79 cm^{-1}). The corresponding absorption line is easily observable in the microwave spectrum of ammonia. In fact, this transition was utilized in the MASER,† the forerunner of the LASER.

---

*The term inversion is used here to refer to a feasible physical phenomenon, which in this case is similar to an umbrella that turns inside-out in the wind. The word has already been used twice in different senses: (i) the inversion of coordinates, *e.g.* the passage from a right-handed system to a left-handed one and (ii) a traditional symmetry operation applied to a molecule with a center of inversion. Accordingly, this term must be used with care!

†Acronym for 'Microwave Amplification by Stimulated Emission of Radiation'.

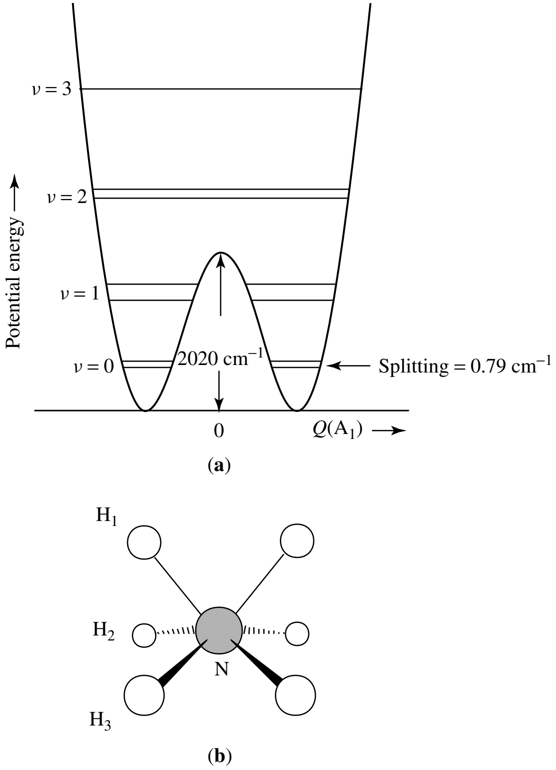


Fig. 5 (**a**) Potential function and energy levels as functions of the normal coordinate $Q(A_1)$; (**b**) Structure of the ammonia molecule in its two equivalent inverted configurations.

The symmetry of the ammonia molecule, as described in Chapter 8 was that of point group $\mathscr{C}_{3v}$. Following that analysis, no possibility of change in handedness is considered to be feasible. It is just this physical criterion that becomes significant in the application of symmetry arguments to nonrigid molecules. Under molecular inversion, the handedness of the structure changes. Thus, symmetry operations such as (132) become feasible. There are, in effect, six of these operations, which, along with those indicated in Table 8-6, define a symmetry group of order 12. It is isomorphic with the group $\mathscr{D}_{3h}$, as given in Table 8-8. The characters are obtained by forming the direct product of the group $\mathscr{C}_s$ with the group $\mathscr{C}_{3v}$. This result serves as the basis for the interpretation of both the microwave and infrared spectra of this molecule. In effect, the infrared spectrum of the gas consists of the replicated P-Q-R features characteristic of a parallel band of a symmetric top.

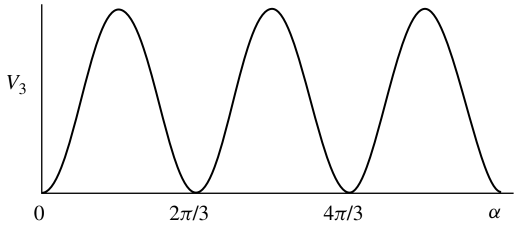


Fig. 6 The potential function for the rotation of a molecule of the ethane type.

9.4.2 Internal rotation

Many molecules are composed of functional groups that can rotate with respect to the rest of the molecule. The classical example is ethane, as the possibility of rotation of one methyl group against the other was recognized long ago. Because the torsional mode does not result in infrared activity, its frequency was estimated from thermodynamic data.

The potential function that governs internal rotation in ethane is represented in Fig. 6. The three equivalent minima correspond to equilibrium positions, that is, three identical molecular structures. The form of this potential function for an internal rotator with three-fold symmetry can be expressed as a Fourier series,

$$V_3 = \frac{1}{2}\sum_{k=1}^{\infty} V_{3k}(1 - cos\ 3k\alpha), \tag{82}$$

where α is the angle of rotation from a given equilibrium position. For most practical purposes the first term ($k = 1$) is sufficient, as it is generally much larger than subsequent terms in the series. It is shown in Section 5.5.6 that the resulting Schrödinger equation is one form of Mathieu's equation. In the limit of a high potential barrier the equation becomes that of a harmonic oscillator and the result is a well-defined torsonal mode of oscillation (problem 23). On the other hand, if the barrier is very low relative to the torsional kinetic energy, the molecule becomes "nonrigid".

The substituted ethanes have been the subject of numerous spectroscopic investigations. A particularly interesting example is provided by 1,2-dichloroethane, which is shown in the *trans* conformation in Fig. 7a, along with its Newman diagram. The barrier to internal rotation is relatively high in this case. Thus, rotamers can be distinguished, depending on the possible equilibrium positions. Figure 7b shows the two other equilibrium configurations. They are known as the *gauche* forms. It should be noted that they are stereoisomers, that is, they are mirror images.

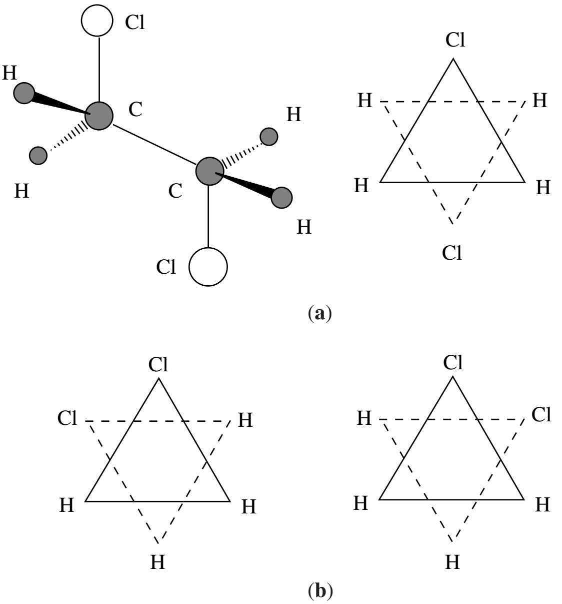


Fig. 7 (**a**) The *trans* configuration of 1,2-dichloroethane. (**b**) The *gauche* isomers of 1,2-dichloroethane.

The symmetry of the *trans* isomer of 1,2-dichloroethane can be described by point group $\mathcal{C}_{2h}$, while the two *gauche* forms are of symmetry $\mathcal{C}_s$. Thus, the spectroscopic selection rules are quite different for these two basic structures. In the *trans* configuration the presence of a center of symmetry precludes the coincidence of fundamental bands in the infrared and Raman spectra. The resulting spectra are much simpler than those of the *gauche* forms. Experimental investigations have shown that the molecule is in the *trans* form in the vapor phase, while it passes to the *gauche* forms under the influence of intermolecular forces in the neat liquid.

Another example of a molecule that exhibits a three-fold internal rotation is dimethyl acetylene, whose structure is represented in Fig. 8. Here the molecular configuration is illustrated in its staggered ($\mathcal{D}_{3d}$) form, although in this particular case the barrier to internal rotation is very low (less than 10 cm^{-1}). The essentially free rotation of the methyl groups allows the cyclic permutation of the hydrogen atoms of a given methyl group, such as [3 1 2 4 5 6] and [2 3 1 4 5 6], to be carried out. However, it can be seen that the operation [1 3 2 4 5 6] is not feasible without the simultaneous inversion of the coordinates, namely, [1 3 2 4 5 6]$^{\#}$, which of course implicates the

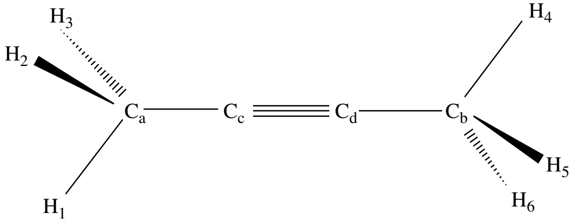


Fig. 8 The dimethyl acetylene molecule.

carbon atoms as well. Clearly, an operation such as [b a c d] that involves the carbon atoms is not feasible. The analysis of this problem from the theoretical point of view leads to a group that is formed by the direct product of two groups of order six that are isomorphic with the group $\mathscr{C}_{3v}$ (see Table 8-6). The resulting group, which is of course of order 36, is not isomorphic with a traditional point group as given in Appendix VIII. It is thus necessary in this case to develop the irreducible representations (or at least their characters) from first principles.

9.4.3 Molecular conformation: The molecular mechanics method

It was shown in the previous section that different relatively stable conformations of a given molecule can result from internal rotation of a particular functional group. The possibility of the existence of various conformers is of extreme importance in many applications. It should be noted, for example, that the biological activity of an organic molecule often depends on its conformation – in particular the relative orientation of a specific functional group. As another example, the thermodynamic properties of, say, an alkane are directly related to the conformation of its carbon skeleton. In this context the industrial importance of *iso*-octane is well-known.

The examples cited above are of molecules which are not strictly speaking nonrigid, although they have more than one well-defined equilibrium configuration. The 1,2-dichloroethane molecule discussed above is a classic example. With the aid of computer programs that have been developed to treat this problem, it has become possible to calculate with confidence the equilibrium conformations of such molecules, as well as the energy differences and the tunneling barriers between them. It is appropriate here to summarize briefly the so-called molecular mechanics method that is currently employed to obtain these results.

Here again a molecule is considered to be a collection of nuclei held together by the electron cloud. This model is based on the Born–Oppenheimer approximation, as described in Section 6.5 and developed in more detail in

Chapter 12. The forces which act on the atoms to maintain them in certain equilibrium positions are associated with changes in bond lengths and angles, and, furthermore in the present application, torsional angles around specific chemical bonds. Once a potential function has been established the so-called steric energy of the molecule can, in principle, be evaluated.

The molecular mechanics method is usually limited to the determination of molecular geometry and thermodynamic quantities. However, it is sometimes employed to estimate vibrational frequencies – at least in those cases in which π electrons are not involved in the determination of the molecular geometry. It should be emphasized that this method, as well as those presented in Chapter 12, are applicable only to isolated molecules, as intermolecular forces are not included in the model.

The application of the molecular mechanics method is carried out in three steps, namely,

(i) the formulation of the potential-energy function,
(ii) the determination of the appropriate parameters in this function, and
(iii) The minimization of the molecular energy.

There exist several programs that employ this procedure. They lead to very similar, although not identical, results.

The steric energy of a molecule depends on the relative positions of its atoms. Starting with a particular geometry the potential energy can be written in the form

$$V = V(\boldsymbol{r}) + V(\boldsymbol{\theta}) + V(\boldsymbol{\tau}) + V(\boldsymbol{d}) + V_{el}, \tag{83}$$

where the first three terms on the right-hand side of Eq. (83) take into account, respectively, the changes in bond lengths, the changes in angles between adjacent bonds and the torsional movements of certain functional groups in the molecule. The fourth term is often necessary to include the forces acting directly between nonbonded atoms. The fifth term is added if there are important electrostatic interactions between polar bonds in the molecule.

The first two terms on the right-hand side of Eq. (83) are usually assumed to be harmonic, as given for example by Eq. (6–74). The third term is often developed in a Fourier series, as given by Eq. (82). The potential function appropriate to the interaction between nonbonded atoms is taken to be of the Lennard-Jones type (Section 6.7.3). In all of these cases the necessary force constants are estimated by comparing the results obtained from a large number of similar molecules. If electrostatic interactions are to be considered, effective atomic charges must be suggested and Coulomb's law applied directly [see Eq. (6–81)].

The final step in the molecular-mechanics calculation of molecular conformation involves the minimization of the energy. Approximations are involved whose importance is not always clear. Usually, all first derivatives with respect to the various internal coordinates are set equal to zero – although these coordinates are often not independent (see Section 10.6). Furthermore, the final conformation obtained depends on the assumed initial structure. Therefore, the method must be applied with care and a certain amount of chemical intuition. In spite of these uncertainties the molecular mechanics method has been employed with considerable success, particularly in the conformational analysis of branched alkanes. For molecules containing hetero-atoms, it can be applied, but with somewhat less confidence.

PROBLEMS

1. Verify Eq. (4).

2. Show that the Eckart conditions [Eqs. (5) and (6)] lead to Eq. (17).

3. Develop Eq. (8) to obtain Eq. (10).

4. Apply Eq. (12) to obtain Eq. (13).

5. Carry out the matrix multiplication indicated in Eq. (17) to verify Table 1.

6. With the use of Eq. (12) derive Eq. (13).

7. Show that the rotational energy of a rigid symmetric top is given by

$$\varepsilon_{rot} = \tfrac{1}{2}[I_a(sin^2\theta\dot{\varphi}^2 + \dot{\theta}^2) + I_c(\dot{\chi} + cos\,\theta\dot{\varphi})^2].$$

8. Derive Eq. (20).

9. Find the inverse of the matrix given in Eq. (18). Ans. See Eq. (21)

10. Verify Eq. (23).

11. Derive Eqs. (24) to (26).

12. Derive Eq. (27).

13. Substitute Eq. (28) in the Schrödinger equation to obtain Eq. (29).

14. Make the indicated substitutions to obtain Eq. (30).

15. Develop the differential equation for the symmetric top [Eq. (32)] and compare it with Eq. (4–154).

16. Derive Eqs. (49) and (50).

17. Derive Eqs. (54) and (55).

18. Verify the matrix $\boldsymbol{B}$ for the water molecule, as given in Eq. (60).

19. Verify the matrix $\boldsymbol{G}$ of Eq. (62).

20. With the use of the $\boldsymbol{G}$ and $\boldsymbol{F}$ matrices [Eqs. (62) and (63)] develop the secular determinant for the water molecule.

21. Verify the matrices $\mathbf{G}$ and $\mathbf{F}$, as given by Eqs. (75) and (76).

22. Develop the character table for the symmetry group $\mathcal{C}_{\mathrm{s}} \otimes \mathcal{C}_{3\mathrm{v}}$.

23. Find the expression for the classical frequency of the torsional vibration of a rigid methyl group (of moment of intertia I) in the limit as $\alpha \to 0$.

$$\text{Ans. } \nu^0 \approx \frac{1}{2\pi}\sqrt{\frac{9V_3}{2I}}$$

10 Probability and Statistics*

The objective of this chapter is to summarize the necessary mathematical concepts employed in the general area of statistical thermodynamics. As molecular systems are composed of a very large number of particles, probability arguments lead to what are essentially exact results. It is useful to begin with some remarks concerning the notion of permutation. It was introduced in Section 8.3, where it was emphasized that it is the permutation of symbols, rather than the permutation of objects, that is basic in the physical application of this mathematical concept. It is with some reservation that in the following sections the word "objects", or sometimes "things", will be used (rather than symbols), as it is thus that the subject is presented in virtually all existing textbooks. This point was illustrated in Section 8.3 by the story of the identical twins.

10.1 PERMUTATIONS

Given n identical (but distinguishable) objects, how can they be placed, say, in n different positions along a line? As they are identical, any one of them, say the first, can be placed with equal probability at any given position. There are then $n - 1$ positions open for the second. Hence, these two objects can be arranged in $n(n - 1)$ different ways. Clearly, the third object can be placed at any of the $n - 2$ remaining places, *etc*. The total number of possible solutions is then given by $n(n - 1)(n - 2) \ldots 2 = n!$. This result expresses the permutation of n things taken n at a time,

$$P(n, n) = n!. \tag{1}$$

In the simplest case, $n = 2$ and the permutation of two identical objects a and b corresponds to the possibilities a,b and b,a or $P(2, 2) = 2$. If a third object, c, is included in all possible ways, the permutations are then $P(3, 3) = 6$.

---

*According to Mark Twain (Samuel Clemens), American author (1835–1910) "There are three kinds of lies: lies, damned lies, and statistics".

The permutation of three identical objects was illustrated in Section 8.3. However, in the application considered there, coordinate systems were used to specify the positions of the particles. It was therefore necessary on the basis of feasibility arguments to include the inversion of coordinates (specified by the symbol #) with those permutations that would otherwise change the handedness of the system. Nevertheless, for the permutation of three particles the order of the group was found to be equal to $3! = 6$.

The application of Eq. (1) can be specifically illustrated by the example of four identical, but distinguishable, objects. The possible permutations can be established with the aid of the following table.

Table 1 Permutation of four distinguishable objects.

a b c d	a c b d	c a b d	b a c d	b c a d	c b a d
a b d c	a c d b	c a d b	b a d c	b c d a	c b d a
a d b c	a d c b	c d a b	b d a c	b d c a	c d b a
d a b c	d a c b	d c a b	d b a c	d b c a	d c b a

Equation (1) is often developed with the use of the following example. In how many ways can n patients occupy n chairs in a doctor's waiting room? The first patient to arrive can choose any one of the n chairs. However, the next patient has only $n - 1$ chairs from which to choose, and so on. Clearly the last patient can take only the single remaining chair. The total number of possibilities is thus given by $n(n-1)(n-2)\ldots 1 = n!$. In this example each possible permutation involves all n patients.

The result obtained above can be generalized by consideration of the case in which there are but r chairs for n patients, where $r < n$. In this case the question is: In how many ways can a given ensemble of n people be seated in the r chairs? By the same argument, the result

$$P(n, r) = n(n-1)(n-2)\ldots(n-r+1) \tag{2}$$

is easily obtained. Clearly, with r chairs filled the following patients cannot be seated. Eq. (2) is usually written in the form

$$P(n, r) = \frac{n!}{(n-r)!}\ , \tag{3}$$

which represents the permutation of n things taken r at a time, or arrangements.

If four distinguishable objects are taken two at a time, the possible arrangements are shown in Table 2. Thus, $P(4, 2) = 12$, a result that can be compared with $P(4, 4) = 24$, as given in Table 1.

Table 2 Arrangement of four distinguishable objects taken two at a time.

a b	a c	a d
b a	b c	b d
c a	c b	c d
d a	d b	d c

10.2 COMBINATIONS

Now consider an example that is closely related to that presented in the previous section. An ensemble of n prospective clients are waiting to be seated in a restaurant. Here, unlike the case of the doctor's waiting room, there is a hostess, whose role is to select the appropriate ensemble of r clients to be seated at a given table. It is now of interest to determine the number of combinations (or selections) of n things r at a time. In this case the order of the members of the ensemble of r persons has not been considered. It is useful, then, to denote the number of combinations by $C(n,r)$, which can be found by the following argument.

The seating of the r patients in the Doctor's office was determined by the specific order in which they arrived. The solution to this problem was found to be the expression given by Eq. (3). However, in the example of the restaurant the ensemble of r persons was chosen by the hostess and then seated in one of $P(r, r) = r!$ possible ways, according to Eq. (1). The total number of ways of selecting and seating the n hungry clients is then given by

$$P(n, r) = C(n, r)P(r, r). \tag{4}$$

Thus, the number of combinations is equal to

$$C(n, r) = \frac{P(n, r)}{P(r, r)} = \frac{n!}{(n - r)!r!} = \binom{n}{r}. \tag{5}$$

The notation employed in Eq. (5) corresponds to that defined in Section 2.10, namely, the binomial coefficients. Thus, in general the binomial coefficients in the polynomial expansion

$$(a + b)^n = \sum_{r=0}^{n} \binom{n}{r} a^{n-r} b^r \tag{6}$$

can be identified with the number of combinations of n "objects" taken r at a time.

As an example of the application of Eq. (5), consider four identical objects taken two at a time. The possible combinations are given in Table 3. As

Table 3 Combinations of four identical objects taken two at a time.

a b	a c	a d
(b a)	b c	b d
(c a)	(c b)	c d
(d a)	(d b)	(d c)

the order is not important, the possibilities in parentheses are not counted, yielding $C(4, 2) = 6$. The division by $r! = 2$ in this example arises from the indistinguishability of the objects in each pair. This result should be compared with those given in Tables 1 and 2.

A closely related problem was considered in the seventeenth century. It is the famous arithmetic triangle developed by Pascal* that is shown in Table 4. It is constructed by writing the number one twice as the first line. The first column is then filled with an infinite sequence of the number one. Subsequently, each value in the table is calculated by taking the sum of the number immediately above and the number to the left of the latter. It is then apparent that the second column is given by $C(n, 1) = n$, the third by $C(n, 2) = n(n - 1)/2$, and in general as given by Eq. (5).

It will be of interest in the following section to evaluate the sum over all given combinations of n objects taken r at a time, namely,

$$\sum_{r=0}^{n} C(n, r) = \sum_{r=0}^{n} \frac{n!}{(n-r)!r!} = \sum_{r=0}^{n} \binom{n}{r}. \tag{7}$$

By setting $a = b = 1$ in Eq. (6), the desired sum is found in the form

$$\sum_{r=0}^{n} C(n, r) = 2^n. \tag{8}$$

Table 4 Pascal's triangle.

1	1						
1	2	1					
1	3	3	1				
1	4	6	4	1			
1	5	10	10	5	1		
1	6	15	20	15	6	1	
1	7	21	35	35	21	7	1
⋮							

---

*Blaise Pascal, French mathematician, philosopher (1623–1662).

10.3 PROBABILITY

The notion of probability is employed in everyday language without, however, a mathematical description. If you say, "it will probably rain today", you do not know the consequence, namely, should you water your garden? You are then dealing with an unknown outcome of a particular event. Probability theory attempts to specify the state of ignorance of the given event. It is applied in many areas of physical science, in particular in quantum chemistry and statistical mechanics. It can be employed with confidence in these areas, as the number of particles and their possible quantum states is very large, say, of the order of Avogadro's number.* Its application in the theory of games is more approximate, as the number of events is usually much more limited. These points will be emphasized in the following examples.

In the development of probability theory, as applied to a system of particles, it is necessary to specify the distribution of particles over the various energy levels of a system. The energy levels may be clearly separated in a quantized system or approach a continuum in the classical limit. The notion of probability is introduced with the aid of the general relation

$$\mathscr{W} = \frac{n}{\mathscr{N}} \tag{9}$$

where n is the number of "favorable" results and $\mathscr{N}$ is the number of possible results.

As a simple example, consider a deck of 52 playing cards. According to Eq. (9), the probability of drawing a heart from the deck is given by

$$\mathscr{W}(\heartsuit) = \frac{13}{52} = \frac{1}{4}, \tag{10}$$

as there are, of course, 13 hearts in the deck. The same result is assumed for the probability of drawing, say, a spade. The probability of drawing either a heart or a spade is then given by the sum

$$\mathscr{W}(\heartsuit) + \mathscr{W}(\spadesuit) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \tag{11}$$

because they are independent results. That the probabilities $\mathscr{W}(\heartsuit)$ and $\mathscr{W}(\spadesuit)$ are equal is an assumption; is it intuitively obvious?

Note, however, that the combined probability of drawing the ace of spades is equal to

$$\mathscr{W}(\mathscr{A})\mathscr{W}(\spadesuit) = \frac{1}{13} \times \frac{1}{4} = \frac{1}{52}. \tag{12}$$

---

*Amedeo di Quaregna Avogadro, Italian chemist (1776–1856).

This result is clear, as there are but four aces in the deck and only one ace of spades.

The probability of a given event is often represented as a function of a random variable, say, x. The random variable can take on various discrete values x_i with probabilities given by $\mathcal{W}(x_i)$. The variable x is then an independent variable that describes a random or stochastic process. The function $\mathcal{W}(x_i)$ in simple examples is discontinuous, although as the number of samples increase, it approaches a denumerable infinity.

Returning now to a game, consider the possible sum of numbers on two dice. Each die has six sides and, assuming that the dice are not loaded, the outcome of the roll of each die has six equally probable possibilities. The probability for each number on a given die is thus equal to 1/6 and the combined probability for the dice is equal to 1/36. The possible values of the sum and the corresponding probabilities are then given in Table 5. These results are plotted in Fig. 1, which represents the probability distribution for this example.

The net result for the probability distribution of the sum of the numbers on two dice is then represented in Fig. 1. The well-known significance of the number seven becomes evident, as it has the greatest probability. As a second roll of the dice is independent of the result of the first, the chances of getting an eleven is only 2/36. However, it is the conditional probability, that is, the

Table 5 The sum of numbers on two dice and the corresponding probability.

Sum	2	3	4	5	6	7	8	9	10	11	12
Probability	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

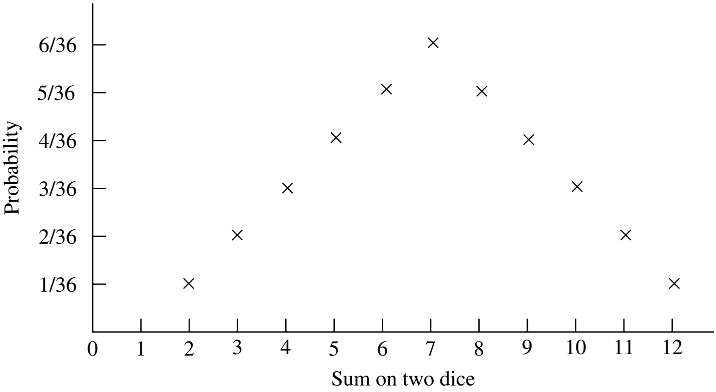


Fig. 1 The probability distribution of the sum of the two numbers on a pair of dice.

chance of obtaining a seven immediately followed by an eleven, that is of interest to the player. The reader is referred to texts on probability theory for discussions of conditional probability, as this subject is usually of less interest in physical applications.

For another game, consider the toss of a coin. If the coin is perfect, the probability of obtaining "heads" (and that of obtaining "tails") is of course equal to one-half. If the coin is tossed five times, the probability of any particular sequence of heads and tails is given by $\left(\frac{1}{2}\right)^5$ because the tosses are independent and the order is unimportant in this case. The number of sequences containing three heads is given by the number of ways of selecting three positions out of five heads, namely, the combination $C(5,3)$. The probability of three heads in five tosses is then given by $\left(\frac{1}{2}\right)^5\ C(5,3)$. This result is expressed in general by the probability function

$$\mathcal{W}(x) = \left(\tfrac{1}{2}\right)^n C(n, x), \tag{13}$$

where x represents the number of heads and n the number of tosses.

In the example presented in the previous paragraph, two results of each event have been considered. In this problem, an event has been repeated with two possible outcomes. One of them is usually referred to as the "success" (p) and the other as "failure" ($q = 1 - p$). These independent events are known as Bernoulli trials, after the Bernoulli brothers.* The general expression for the probability is then given by

$$\mathcal{W}(x) = p^x q^{n-x} C(n, x). \tag{14}$$

For a coin $p = q = \frac{1}{2}$, and Eq. (14) reduces to Eq. (13).

In the examples described above the resulting probability distributions were discontinuous functions. However, it is not difficult to imagine cases in which the distributions become continuous in the limit of an infinite – or at least a very large – number of trials. Such is the case in the application of statistical arguments to problems in thermodynamics, as outlined in Section 10.5.

10.4 STIRLING'S APPROXIMATION

Factorial numbers enter often in the equations developed in the previous sections. They are not very conveniently evaluated in the applications considered below. However, there exists an approximation due to Stirling† that is

---

*The Bernoulli family of Belgian–Swiss mathematicians, of whom Daniel (1700–1782) is probably the best known.

†James Stirling, British mathematician (1692–1770).

very useful in the evaluation of factorials of large numbers. It can be derived in several ways. For example, the Gamma function was defined by Eq. (5-143) in the form

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1}\, \mathrm{d}t. \tag{15}$$

If z is replaced by $n + 1$, a positive integer, Eq. (15) becomes

$$\begin{aligned}\Gamma(n+1) = n! &= \int_0^\infty e^{-t} t^n\, \mathrm{d}t \\ &= \int_0^\infty e^{n\, ln\, t - t}\, \mathrm{d}t.\end{aligned} \tag{16}$$

With the substitution $t = n + y\sqrt{n}$, and hence $\mathrm{d}t = \sqrt{n}\, \mathrm{d}y$, Eq. (16) becomes

$$n! = \int_{-\sqrt{n}}^\infty e^{n\, ln\, (n + y\sqrt{n}) - n - y\sqrt{n}} \sqrt{n}\, \mathrm{d}y. \tag{17}$$

It should be noted that at $t = 0$, $y = -\sqrt{n}$. The logarithm appearing in the exponent in Eq. (17) can be developed in the form

$$ln(n + y\sqrt{n}) = ln\, n + ln\left(1 + \frac{y}{\sqrt{n}}\right) = ln\, n + \frac{y}{\sqrt{n}} - \frac{y^2}{2n} + \ldots \tag{18}$$

(see problem 1). Substitution of Eq. (18) into Eq. (17) leads to the approximation

$$\begin{aligned} n! &\approx \int_{-\sqrt{n}}^\infty e^{n\, ln\, n + y\sqrt{n} - y^2/2 - n - y\sqrt{n}} \sqrt{n}\, \mathrm{d}y \\ &\approx e^{n\, ln\, n - n} \sqrt{n} \int_{-\sqrt{n}}^\infty e^{-y^2/2}\, \mathrm{d}y. \end{aligned} \tag{19}$$

The integral in Eq. (19) can be evaluated in terms of the error function, but for practical purposes it can be replaced by $\sqrt{2\pi}$ in the limit of large values of n (see Section 3.4.5). Then,

$$n! \approx n^n e^{-n} \sqrt{2\pi n}. \tag{20}$$

In applications in statistical mechanics it is the logarithm of large factorials that is of primary interest. The logarithm of Eq. (20) can be written in the limit of large values of n as

$$ln\, n! = \left(n + \tfrac{1}{2}\right) ln\, n - n + \tfrac{1}{2}\, ln\, 2\pi. \tag{21}$$

$$\approx n\, ln\, n - n. \tag{22}$$

Equation (22) is employed repeatedly in the applications illustrated in the following sections. It should be noted that this relation can also be obtained by evaluation of $\int_1^n \ln x \mathrm{d}x$ (see problem 2).

10.5 STATISTICAL MECHANICS

In physical chemistry the most important application of the probability arguments developed above is in the area of statistical mechanics, and in particular, in statistical thermodynamics. This subject supplies the basic connection between a microscopic model of a system and its macroscopic description. The latter point of view is of course based on the results of experimental measurements (necessarily carried out in each experiment on a very large number of particles) which provide the basis of classical thermodynamics. With the aid of a simple example, an effort will now be made to establish a connection between the microscopic and macroscopic points of view.

Consider a system composed of n identical, but distinguishable, particles. The distinguishability of the particles may result, for example, from their positions in space, *e.g.* their coordinates. It is useful in this simplified model to assume, furthermore, that the energy of interaction between the particles can be neglected. Thus, n_1 particles in the ensemble have energy ε_1 *etc.*, and n_i the energy ε_i. The number of ways of choosing n_1 out of the ensemble of n particles is given by Eq. (5), with $r = n_1$, namely,

$$C(n, n_1) = \frac{n!}{(n - n_1)!n_1!} \,. \tag{23}$$

There remain but $n - n_1$ particles from which to choose the number n_2. This result is easily generalized to obtain the probability distribution involving all n particles in the form

$$\mathscr{W} = \frac{n!}{n_1!n_2!n_3!\cdots} = n!\prod_i \frac{1}{n_i!} \,. \tag{24}$$

It should be noted that each factorial appearing in the denominator of Eq. (24) arises from the permutation of the n_i identical particles of energy ε_i.

In the derivation of Eq. (24) the possibility of degeneracy was not considered. However, if there are g_i states with the same energy ε_i, this result can be generalized. In effect, the first particle among the n_i can be placed in g_i ways among the states of energy ε_i. As there is no limitation here on the number of particles in each state, there are $g_i^{n_i}$ ways to place the n_i particles. Thus, the more general form of Eq. (24) becomes

$$\mathscr{W} = n! \prod_i \frac{g_i^{n_i}}{n_i!} \, . \tag{25}$$

With the use of Stirling's approximation, as given by Eq. (22), the logarithm of Eq. (25) can be written as

$$ln\, \mathscr{W} = \sum_i n_i \, ln\, g_i - \sum (n_i \, ln\, n_i - n_i) + n \, ln\, n - n. \tag{26}$$

It should be pointed out that the n particles have been assumed here to be distinguishable. However, in an ideal gas the molecules cannot be distinguished, as their positions are random in space. Therefore, to be applied to the case of an ideal gas, Eqs. (24) and (25) should be divided by $n!$. Equation (26) becomes simply,

$$ln\, \mathscr{W} \approx \sum_i n_i \, ln\, g_i - \sum (n_i \, ln\, n_i - n_i). \tag{27}$$

It might seem that the most probable distribution could be found by setting the differential $\delta \mathscr{W}$ equal to zero. Or, as $\mathscr{W}$ and $ln\, \mathscr{W}$ are maximum at the same point (if $\mathscr{W} \neq 0$), the differential of Eq. (27) can also be considered. The needed result is given by (see problem 3)

$$\delta \, ln\, \mathscr{W} = ln\, \frac{g_i}{n_i} \delta n_i. \tag{28}$$

Note, however, that the condition that Eq. (28) vanishes is subject to two constraints, namely,

$$n = \sum_i n_i \tag{29}$$

and

$$E = \sum_i \varepsilon_i n_i \tag{30}$$

which must be imposed to obtain a statistical description of an ideal gas. Equations (29) and (30) express the conservation of the number of particles (and, hence, mass) and the energy, respectively. The problem is then to maximize $ln\, \mathscr{W}$ and at the same time assure the conditions given by Eqs. (29) and (30).

10.6 THE LAGRANGE MULTIPLIERS

Before the results obtained in the previous section can be applied, it is necessary to describe briefly the method of underdetermined multipliers. Given a function $f(x_1, x_2, \ldots, x_\mathrm{k})$ of variables x_i, for which it is desired to find stationary values, the chain rule leads to the expression

$$\mathrm{d}f = \left(\frac{\partial f}{\partial x_1}\right)\mathrm{d}x_1 + \left(\frac{\partial f}{\partial x_2}\right)\mathrm{d}x_2 + \ldots + \left(\frac{\partial f}{\partial x_k}\right)\mathrm{d}x_k, \tag{31}$$

as given in Section 2.12. If the variables $x_1, x_2, \ldots, x_\mathrm{k}$ are independent, the stationery values of $f(x)$ can be found by setting each partial derivative in Eq. (31) equal to zero. However, if there are constraints on the system, that is, if the variables x_i are not independent, the problem becomes more complicated.

Consider the case in which there are two implicit relations among the variables, *e.g.*

$$g(x_1, x_2, \ldots, x_k) = 0 \tag{32}$$

and

$$h(x_1, x_2, \ldots, x_k) = 0. \tag{33}$$

Expressing Eqs. (32) and (33) with the use of the chain rule leads to three relations which must be simultaneously satisfied, *viz.*

$$\sum_i \frac{\partial f}{\partial x_i}\,\mathrm{d}x_i = 0, \tag{34}$$

$$\sum_i \frac{\partial g}{\partial x_i}\,\mathrm{d}x_i = 0 \tag{35}$$

and

$$\sum_i \frac{\partial h}{\partial x_i}\,\mathrm{d}x_i = 0. \tag{36}$$

If Eqs. (35) and (36) are each multiplied by an arbitrary constant, say $-\alpha$ and $-\beta$, respectively, and added, the result is given by

$$\sum_i \left(\frac{\partial f}{\partial x_i} - \alpha\frac{\partial g}{\partial x_i} - \beta\frac{\partial h}{\partial x_i}\right)\mathrm{d}x_i = 0. \tag{37}$$

As the various differentials $\mathrm{d}x_i$ are not identically zero, the expression in parentheses must be, leading to the expression

$$\frac{\partial f}{\partial x_i} = \alpha\frac{\partial g}{\partial x_i} + \beta\frac{\partial h}{\partial x_i}. \tag{38}$$

In this example α and β are the Lagrange multipliers, which can be determined from the expressions for the constraints, although it is not always necessary to do so. Clearly, this approach can be generalized and the number of multipliers is equal to the number of constraints on the system.

10.7 THE PARTITION FUNCTION

The statistical description of an ideal gas, a system of n identical, indistinguishable (nonlocalized) particles, can now be obtained from Eqs. (28) to (30). The differentials of Eqs. (29) and (30) are given by

$$\sum_i \delta n_i = 0 \tag{39}$$

$$\sum_i \varepsilon_i \delta n_i = 0, \tag{40}$$

respectively. The problem of maximizing $ln\, \mathscr{W}$ subject to these constraints can be expressed by

$$\sum_i \left(ln\, \frac{g_i}{n_i} - \alpha - \beta\varepsilon_i \right) \delta n_i = 0, \tag{41}$$

where α and β are Lagrange multipliers. As the quantity in parentheses must vanish in each term, the number of particles with energy ε_i is given by

$$n_i = g_i e^{-\alpha-\beta\varepsilon_i}. \tag{42}$$

The evaluation of the Lagrange multipliers requires additional information.

From Eqs. (42) and (29) the fraction of particles in energy level i becomes

$$\frac{n_i}{n} = \frac{g_i e^{-\beta\varepsilon_i}}{\sum_i g_i e^{-\beta\varepsilon_i}}. \tag{43}$$

The denominator of Eq. (43), $\sum_i g_i e^{-\beta\varepsilon_i} \equiv \mathscr{Z}$, is known as the partition function (French: *fonction de partition*) or state sum (German: *Zustandssumme*). It will be shown later that $\beta = 1/kT$, where T is the absolute temperature and k is the Boltzmann constant.* The partition function is thus a function of the absolute temperature.

---

*In classical thermodynamics the (inexact) differential change in heat, δq, is related to the (exact) differential change in entropy, dS, by $\mathrm{d}S = k\beta\delta q = T^{-1}\delta q$. Thus, $k\beta = 1/T$ is an integrating factor for the heat change (see Section 3.5).

The substitution of Eq. (43) in Eq. (30) yields the expression for the internal energy of the system,

$$\begin{aligned} E &= \frac{n}{\mathcal{Z}} \sum_i \varepsilon_i g_i e^{-\varepsilon_i/kT} \\ &= \frac{nkT^2}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial T}. \end{aligned} \tag{44}$$

The energy per mole of gas is then equal to

$$\tilde{E} = RT^2 \frac{\partial \, ln \, \mathcal{Z}}{\partial T}, \tag{45}$$

where the tilde indicates a molar quantity. The heat capacity per mole is then given by

$$\tilde{C}_V \equiv \left(\frac{\partial \tilde{E}}{\partial T} \right)_V = R \frac{\partial}{\partial T} \left(T^2 \frac{\partial \, ln \, \mathcal{Z}}{\partial T} \right). \tag{46}$$

From Eqs. (45) and (46) it is apparent that the calculation of the energy and heat capacity of a system depends on the evaluation of the partition function as a function of temperature. In the more general case of molecules with an internal structure, the energy distributions of the various degrees of freedom must be determined. This problem is outlined briefly in the following section.

10.8 MOLECULAR ENERGIES

It was argued in the introduction to Chapter 9 that the energy of a molecule can be written in a first approximation as the sum

$$\varepsilon_i \approx \varepsilon_i(\text{electrons}) + \varepsilon_i(\text{translation}) + \varepsilon_i(\text{rotation}) + \varepsilon_i(\text{vibration}), \tag{47}$$

where the last three terms refer to displacement of the nuclei and all interaction terms have been neglected. For an ensemble of independent molecules, as in an ideal gas, it is apparent from Eq. (45) that the partition function can be written as the product of the appropriate partition functions. Thus,

$$\mathcal{Z} = \mathcal{Z}(\text{translation}) \cdot \mathcal{Z}(\text{rotation}) \cdot \mathcal{Z}(\text{vibration}), \tag{48}$$

where $\mathcal{Z}(\text{electrons}) = 1$, if the molecules are in their (nondegenerate) ground state. However, at very high temperature the population of excited electronic states cannot be neglected and the electronic partition function must be evaluated.

10.8.1 Translation

The displacement of the center of mass of a molecule gives rise to the external or translation energy of the molecule. It corresponds to the energy associated with the movement in three-dimensional space of a particle whose mass is equal to that of the molecule. In each of the three directions in space the problem can be treated as a particle in a box, as presented in Section 5.4.1. It was shown that the energy is given by

$$\varepsilon_i = \frac{h^2 n_x^2}{8ma^2}\,, \tag{49}$$

where the notation has now been modified so that $n_x = 1, 2, 3, \ldots$ is the quantum number associated with the translation in the x direction and a is the corresponding dimension of the box. With the use of the definition of the partition function, as given below Eq. (43), the partition function associated with the x degrees of translational freedom of the molecules is given by

$$\mathcal{Z}^{(x)} = \sum_{\mathrm{N}_x} e^{-\frac{h^2 n_x^2}{8ma^2kT}}\,, \tag{50}$$

where β has been replaced by $1/kT$, as before, and the system is nondegenerate ($gn_x = 1$ for all values of n_x). From Eq. (49) it is apparent that the separation between successive energy levels decreases as the quantity ma^2 increases. As the dimension a is very large on a molecular scale, for virtually all molecules the energy distribution given by Eq. (49) approaches a continuum. In other words, the deBroglie wavelength of the particle is very short compared with a, the dimension of the container. At moderate temperatures almost all of the molecules are to be found in the lower levels. Thus, to a good approximation, the sum in Eq. (50) can be replaced by an integral over the energy levels, namely,

$$\begin{aligned} \mathcal{Z}^{(x)} &\approx \int_0^\infty e^{-\frac{h^2 n_x^2}{8ma^2kT}}\,\mathrm{d}n_x \\ &= \frac{a\sqrt{2\pi mkT}}{h}. \end{aligned} \tag{51}$$

Obviously, the above derivation can be repeated for the other two Cartesian directions. As the energies are additive, the partition function for the three-dimensional translation of the molecule can be written as a product, *viz.*

$$\begin{aligned} \mathcal{Z}(\text{translation}) &= \mathcal{Z}^{(x)} \cdot \mathcal{Z}^{(y)} \cdot \mathcal{Z}^{(z)} \\ &= \frac{abc(2\pi mkT)^{3/2}}{h^3} = \frac{V}{h^3}(2\pi mkT)^{3/2}, \end{aligned} \tag{52}$$

where $V = abc$ is the volume of the rectangular container. Equation (52) leads to

$$ln\, \mathcal{Z}(\text{translation}) = \tfrac{3}{2}\, ln\, T + \text{constant} \tag{53}$$

and

$$\frac{\partial\, ln\, \mathcal{Z}(\text{translation})}{\partial T} = \frac{3}{2T}. \tag{54}$$

Then, from Eqs. (45) and (46)

$$E(\text{translation}) = \tfrac{3}{2} nkT \tag{55}$$

and the energy per mole is given by

$$\tilde{E}(\text{translation}) = \tfrac{3}{2} RT. \tag{56}$$

This result is identical to that obtained in elementary kinetic theory. There is, in effect a contribution of $\frac{1}{2}RT$ to the energy per mole for each of the three directions in space. Furthermore, this result allows the identification $\beta = 1/kT$, as suggested earlier.

10.8.2 Rotation

It was shown in the previous chapter that the Schrödinger equation for molecular rotation depends on the type of rotator, as defined in Section 9.2.2. For linear molecules and, hence, diatomics, the energy is given by Eq. (9-40),

$$\varepsilon_J = \frac{h^2}{8\pi^2 I} J(J+1), \tag{57}$$

with $J = 0, 1, 2, \ldots$. As the degree of degeneracy is equal to $g_J = 2J + 1$, the rotational partition function can be written as

$$\mathcal{Z}(\text{rotation}) = \sum_J (2J+1) e^{-h^2 J(J+1)/8\pi^2 I kT}. \tag{58}$$

If the molecule does not have a center of symmetry, the sum appearing in Eq. (58) extends over all values of the quantum number J. It should be noted, however, that in the case of centrosymmetric molecules the role of nuclear spins must be considered.

For linear molecules that lack a center of symmetry, Eq. (58) is applicable and at temperatures significantly greater than the "rotational temperature", $\theta \equiv h^2/8\pi^2 Ik$, the sum in Eq. (58) can be replaced by an integral. This operation is equivalent to the passage from the quantum-mechanical to the classical description of the rotational motion of the molecule. It yields

$$\mathscr{Z}(\text{rotation}) \approx \frac{8\pi^2 IkT}{h^2} = \frac{T}{\theta} \tag{59}$$

for $T \gg \theta$ (problem 5). The derivative of the logarithm of Eq. (59) is then given by

$$\frac{\partial\, ln\, \mathscr{Z}(\text{rotation})}{\partial T} = \frac{1}{T}\,, \tag{60}$$

which leads to RT as the rotational energy per mole of a gas composed of linear (noncentrosymmetric) molecules. This result corresponds to a contribution of $\frac{1}{2}RT$ for each of the two degrees of rotational freedom of a linear molecule.

If a linear molecule contains a center of symmetry, the quantum number J is restricted to either odd or even values. Thus, the sum in Eq. (58) is composed of two parts that approach the same value as J increases, *viz.*

$$\sum_{J\,\text{even}} (2J+1)e^{-J(J+1)\theta/T} \approx \sum_{J\,\text{odd}} (2J+1)e^{-J(J+1)\theta/T} \approx \frac{T}{2\theta}. \tag{61}$$

In effect, the division by two is the result of the molecular symmetry, as specified by the character table for the group $\mathscr{D}_{\infty\text{h}}$. In general it is useful to define a symmetry number σ (= 2 in this case), as shown below. The well-known example of the importance of nuclear spin is that of *ortho*- and *para*-hydrogen (see Section 10.9.5).

Nonlinear polyatomic molecules require further consideration, depending on their classification, as given in Section 9.2.2. In the classical, high-temperature limit, the rotational partition function for a nonlinear molecule is given by

$$\mathscr{Z}(\text{rotation}) \approx \frac{\sqrt{\pi}}{\sigma h^3}\sqrt{I_a I_b I_c}(8\pi^2 kT)^{3/2}, \tag{62}$$

where $I_a = I_b = I_c$ for spherical rotators and $I_a = I_b$ for symmetric rotators. The symmetry number depends on the structure of the molecule. For example, a molecule such as H_2O, which belongs to point group $\mathscr{C}_{2\text{v}}$, has a two-fold axis of symmetry, leading to $\sigma = 2$. For ammonia ($\mathscr{C}_{3\text{v}}$), $\sigma = 3$, while for methane ($\mathscr{T}_\text{d}$) and benzene ($\mathscr{C}_{6\text{v}}$), $\sigma = 4 \times 3 = 12$ (4 three-fold axes) and $\sigma = 6 \times 2 = 12$ (6 two-fold axes), respectively. In all cases Eq. (62) yields $\frac{3}{2}RT$ for the energy per mole of a gas composed of nonlinear molecules.

10.8.3 Vibration

The vibrational energy of a diatomic molecule is given in the harmonic approximation by

$$\varepsilon_v(\text{vibration}) = h\nu^0\left(v + \tfrac{1}{2}\right), \tag{63}$$

where $v = 0, 1, 2, \ldots$ is the vibrational quantum number and ν^0 is the classical frequency of vibration (see Sections 5.4.4 and 6.5.2). The corresponding partition function is then equal to

$$\mathscr{Z}(\text{vibration}) = e^{-h\nu^0/2kT} \sum_v e^{-(h\nu^0/kT)v}. \tag{64}$$

The summation appearing in Eq. (64) can be written in the form

$$\begin{aligned}\sum_v e^{-(h\nu^0/kT)v} &= \sum_v (e^{-h\nu^0/kT})^v \\ &= \sum_v \varsigma^v = 1 + \varsigma + \varsigma^2 + \ldots = \frac{1}{1-\varsigma},\end{aligned} \tag{65}$$

where $\varsigma = e^{-h\nu^0/kT}$. In Eq. (65) ς^v has been developed as a geometric series (see problem 6). Thus,

$$ln\, \mathscr{Z}(\text{vibration}) = -\frac{h\nu^0}{2kT} - ln(1 - e^{-h\nu^0/kT}) \tag{66}$$

and from Eq. (45) the energy per mole is given by

$$\tilde{E} = N_0(\tfrac{1}{2}h\nu^0) + \frac{N_0 h\nu^0 e^{-h\nu^0/kT}}{e^{-h\nu^0/kT} - 1}. \tag{67}$$

The first term on the right-hand side of Eq. (67) is known as the zero-point energy, as it remains even at $T = 0$. It does not, however, contribute to the heat capacity, as Eq. (46) leads to

$$\tilde{C}_V(\text{vibration}) = \left(\frac{\partial \tilde{E}}{\partial T}\right)_V = R\left(\frac{h\nu^0}{kT}\right)^2 \frac{e^{h\nu^0/kT}}{(e^{h\nu^0/kT} - 1)^2} \tag{68}$$

(problem 7). The function

$$f(x) = \frac{x^2 e^x}{(e^x - 1)^2} \tag{69}$$

appearing in Eq. (68) with $x \equiv h\nu^0/kT$ is known as the Einstein function.* In the limit as $kT \gg h\nu^0$, $f(x) \to 1$ (see problem 8), and the heat capacity becomes equal to R in the high-temperature limit. On the other hand the heat capacity is equal to zero at zero Kelvin.

The results obtained above for a diatomic molecule can be generalized for polyatomic molecules. Each of the $3N-6$ normal modes of vibration (or $3N-5$ for linear molecules) will contribute an energy given by an expression analogous to Eq. (63), namely,

$$\varepsilon_{v_k} = h\nu_k^0\left(v_k + \tfrac{1}{2}\right), \tag{70}$$

where v_k is the vibrational quantum number for the k^{th} mode and ν_k^0 is its frequency. Thus, there will be a partition function $\mathcal{Z}_{v_k}$ for each mode such that

$$\mathcal{Z}(\text{vibration}) = \prod_{v_k=1}^{3N-6} \mathcal{Z}_{v_k}. \tag{71}$$

The various contributions to the energy of a molecule were specified in Eq. (47). However, the fact that the electronic partition function was assumed to be equal to one should not be overlooked. In effect, the electronic energy was assumed to be equal to zero, that is, that the molecule remains in its ground electronic state. In the application of statistical mechanics to high-temperature systems this approximation is not appropriate. In particular, in the analysis of plasmas the electronic contribution to the energy, and thus to the partition function, must be included.

10.9 QUANTUM STATISTICS

10.9.1 The indistinguishability of identical particles

A basic principle in quantum mechanics is the indistinguishability of particles. Thus, as indicated in Section 10.5, two particles of the same type in an ideal gas are characterized by a wavefunction, say $\psi(r_1, \theta_1, \varphi_1; r_2, \theta_2, \varphi_2)$, where r, θ, φ are the usual spherical polar coordinates. If for simplicity this wavefunction is written as $\psi(1, 2)$, the permutation of the coordinates of the two identical particles can be represented by

$$\hat{P}\psi(1,2) = \psi(2,1), \tag{72}$$

---

*Albert Einstein, German–American theoretical physicist (1879–1955).

where $\hat{P}$ is the permutation operator (see Section 10.1). Clearly, if this operation is carried twice in succession the result is to re-establish the original description of the system, namely,

$$\hat{P}\hat{P}\psi(1,2) = \hat{P}\psi(2,1) = \psi(1,2). \tag{73}$$

It is useful at this point to define the eigenvalue λ of this operator by the relation

$$\hat{P}\psi(1,2) = \lambda\psi(1,2). \tag{74}$$

The successive application of this operation yields

$$\hat{P}\hat{P}\psi(1,2) = \lambda\hat{P}\psi(1,2) = \lambda^2\psi(1,2). \tag{75}$$

Comparison of Eqs. (73) and (75) shows that $\lambda^2 = 1$; thus, $\lambda = \pm 1$.

The result of the above argument is that the wavefunctions can be classified as either:

(i) symmetric, with no change in sign under the permutation; $\psi(1,2) = \psi(2,1)$

or

(ii) antisymmetric, with change in sign, *viz.* $\psi(1,2) = -\psi(2,1)$.

The conclusion is then that the wavefunction representing a system composed of indistinguishable particles must be either symmetric or antisymmetric under the permutation operation. On purely physical grounds, this result is apparent, as the probability density must be independent of the permutation of indistinguishable particles; or $|\psi(1,2)|^2 = |\psi(2,1)|^2$.

Most particles of interest to physicists and chemists are found to be antisymmetric under permutation. They include electrons, protons and neutrons, as well as positrons and other “antiparticles”. These particles, which are known as Fermions,* all have spins of one-half. The relation between the permutation symmetry and the value of the spin has been established by experiment and, in the case of the electron, by application of relativistic quantum theory.

10.9.2 The exclusion principle

The symmetry of particles [(i) or (ii), above] determines the rules under which they occupy the various quantum states of a system. In the case of Fermions, for example, no more than one particle can occupy a given quantum state.

---

*Enrico Fermi, Italian physicist (1901–1954).

This statement, as applied to electrons, provides the basis for the exclusion principle of Pauli.*

The method of assuring the antisymmetry of a system of electrons, as for example in a polyelectronic atom, is to construct what is often called the Slater determinant.† If the N electrons are numbered $1, 2, 3, \dots$ and each can occupy a state $a, b, c, \dots$, the determinant

$$\psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_a(1) & \psi_b(1) & \psi_c(1) & \cdots & \psi_N(1) \\ \psi_a(2) & \psi_b(2) & \psi_c(2) & & \vdots \\ \psi_a(3) & \psi_b(3) & \psi_c(3) & & \\ \vdots & & & & \\ \psi_a(N) & \cdots & & & \psi_N(N) \end{vmatrix} \tag{76}$$

provides a description of the system. If any two electrons are permuted, the result is to exchange the corresponding two rows in the determinant. The determinant thus changes sign and the wavefunction ψ is necessarily antisymmetric, as required. Furthermore, if any two electrons are found in the same state, *e.g.* $\psi_a = \psi_b$, two columns are identical and the determinant vanishes. Thus, the system does not exist and this possibility is excluded.

The properties of the Slater determinant demonstrate immediately the Pauli exclusion principle, as usually taught. It reads: No two electrons can have all four quantum numbers equal, that is to say that they cannot occupy the same quantum state. It is the direct result of the more general argument that the wavefunction must be antisymmetric under the permutation of any pair of (identical and indistinguishable) electrons.

10.9.3 Fermi–Dirac‡ statistics

These restrictions, imposed above on electrons, apply equally to all particles that are represented by antisymmetric wavefunctions, the so-called Fermions. The condition that no more than one particle can occupy a given quantum state leads immediately to the expression for the number of possible combinations. If $C(n_i, g_i)$ is the number of combinations that can be made with g_i particles taken n_i at a time,

$$C(n_i, g_i) = \frac{g_i!}{n_i!(g_i - n_i)!} \ , \tag{77}$$

as they are indistinguishable and the order in which they are taken is immaterial. For an ensemble of n particles the total number of combinations is

---

*Wolfgang Pauli, Austro-American physicist (1900–1958).

†John C. Slater, American theoretical physicist (1900–1976).

‡P. A. M. Dirac, British theoretical physicist (1902–1984).

given by

$$\mathcal{W} = \prod_i \frac{g_i!}{n_i!(g_i - n_i)!}\ , \tag{78}$$

the product of the number of possible combinations for each energy level. This result is appropriate to Fermi–Dirac statistics, which is applicable to antisymmetric particles.

It is the logarithm of the number of combinations that enters in the appropriate statistics; thus it is the quantity

$$ln\ \mathcal{W} = \sum_i [ln\ g_i! - ln\ n_i! - ln(g_i - n_i)!] \tag{79}$$

that is of interest. Applying Stirling's approximation, Eq. (79) reduces to

$$ln\ \mathcal{W} = \sum_i \left[n_i\ ln\left(\frac{g_i}{n_i} - 1\right) - g_i\ ln\left(1 - \frac{n_i}{g_i}\right)\right]. \tag{80}$$

As a simple illustration of Eq. (77), consider a system composed of three particles of the same energy that can occupy four quantum states. The case of Fermi–Dirac statistics is shown in Fig. 2a.

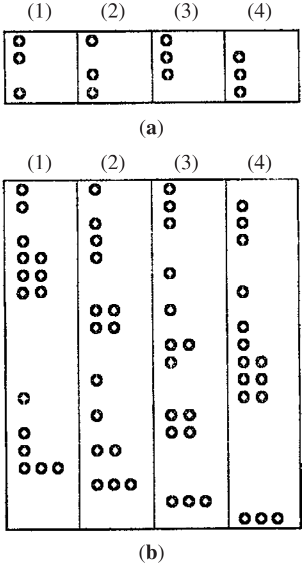


Fig. 2 The distributions of three particles in four quantum states (1) . . . (4) following (**a**) Fermi–Dirac statistics (**b**) Bose–Einstein statistics.

10.9.4 Bose*–Einstein statistics

Under the conditions determined by the symmetric permutation of identical particles, any number of particles can be placed in each quantum state. To distribute n_i of a given energy in g_i quantum states, consider the following model. To create g_i little boxes ("all in a row") only $g_i - 1$ barriers separating the boxes are necessary. There are then n_i particles and $g_i - 1$ barriers. Therefore, there are $n_i + g_i - 1$ Bosons that can be permuted and $(n_i + g_i - 1)!$ possible permutations. However, as the particles are indistinguishable, the elements of each ensemble of n_i can be permuted, leading to $n_i!$ possibilities. Clearly, the indistinguishability of the barriers corresponds to $(g_i - 1)!$ permutations. The number of combinations is then given by

$$C(n_i, n_i + g_i - 1) = \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}. \tag{81}$$

Again, by analogy with Eq. (78), for an ensemble of n_i Bosons the number of combinations is given by

$$\mathscr{W} = \prod_i \frac{(n_i + g_i - 1)}{n_i!(g_i - 1)!}. \tag{82}$$

In the applications of this result in statistical mechanics $g_i \gg 1$ and Eq. (81) reduces to

$$\mathscr{W} = \prod_i \frac{(n_i + g_i)!}{n_i!g_i!}. \tag{83}$$

Here, again, it is the logarithm of this function that is of interest. It is given by

$$\mathit{ln}\, W = \sum_i \left[n_i \,\mathit{ln} \left(\frac{g_i}{n_i} + 1 \right) + g_i \,\mathit{ln} \left(1 + \frac{n_i}{g_i} \right) \right]. \tag{84}$$

Equations (80) and (84) can be written as a single expression, *viz.*

$$\mathit{ln}\, W = \sum_i \left[n_i \,\mathit{ln} \left(\frac{g_i}{n_i} - \iota \right) - \iota g_i \,\mathit{ln} \left(1 - \iota \frac{n_i}{g_i} \right) \right], \tag{85}$$

where ι is equal to ± 1.

With the use of the method developed in Section 10.7, Eq. (85) leads to

$$n_i = \frac{g_i e^{-\alpha - \beta\varepsilon_i}}{1 + \iota e^{-\alpha - \beta\varepsilon_i}}, \tag{86}$$

---

*Satyendranath Bose, Indian physicist (1894–1974).

where α and β are the Lagrange multipliers, as before. If the number of particles of a given energy is much smaller the number of quantum states that they can occupy, $e^{-\alpha-\beta\varepsilon_i} \ll 1$ and Eq. (86) becomes identical to Eq. (42).

In summary, Eq. (86) is a general expression for the number of particles in a given quantum state. If $\iota = \pm 1$, this result is appropriate to Fermi–Dirac statistics, or to Bose–Einstein statistics, respectively. However, if ι is equated to zero, the result corresponds to the Maxwell–Boltzmann distribution. In many cases the last is a good approximation to quantum systems, which is, furthermore, a correct description of classical ones – those in which the energy levels form a continuum. From these results the partition functions can be calculated, leading to expressions for the various thermodynamic functions for a given system. In many cases these values, as obtained from spectroscopic observations, are more accurate than those obtained by direct thermodynamic measurements.

10.10 *ORTHO*- AND *PARA*-HYDROGEN

In Section 10.8 it was assumed that the total energy of a molecule can be written as the sum given in Eq. (47). Accordingly, the partition function was written as a product of the corresponding partition functions. However, a particular problem arises in the case in which the molecule possesses a center of inversion. The simplest case is of course a homonuclear diatomic molecule. The inversion operation with respect to the center of symmetry results in the exchange of the identity of the two nuclei. The effect of this permutation on the wavefunction for the nuclear spins is either symmetric or antisymmetric. If this wavefunction is symmetric, the molecule is called *ortho* and if it is antisymmetric, it is called *para*. It can be shown that the number of independent spin wavefunctions (spin degeneracy or statistical weight) is given by $g_i(ortho) = (i+1)(2i+1)$ if the wavefunction is symmetric under the permutation and $g_i(para) = i(2i+1)$ if it is antisymmetric. Here, i is the spin of one of the identical nuclei. It has integer values or half-integer values, depending on the nature of the nucleus. In the former case, it is called a Boson, while in the latter it is known as a Fermion. The proton, for example, has a spin of one-half and is thus a Fermion (see Section 10.9.1).

According to the argument presented above, any molecule must be described by wavefunctions that are antisymmetric with respect to the exchange of any two identical particles. For a homonuclear diatomic molecule, for example, the possibility of permutation of the two identical nuclei must be considered. Although both the translational and vibrational wavefunctions are symmetric under such a permutation, the parity of the rotational wavefunction depends on the value of J, the rotational quantum number. It can be shown that the wavefunction is symmetric if J is even and antisymmetric if J is odd. The overall symmetry of a homonuclear diatomic molecule is determined by the symmetry of the product of the rotational and nuclear-spin wavefunctions. Thus, for the *ortho* species, the rotational wavefunction is antisymmetric, corresponding to odd values of J. On the other hand, only even values of J are involved in the partition function for the *para* molecules. The specific forms of the partition functions are then

(i) *Ortho* molecules:

$$\mathcal{Z}(\text{nuclear})\mathcal{Z}(\text{rotation}) = (i+1)(2i+1)\sum_{J\text{odd}}(2J+1)e^{-J(J+1)\theta/T} \tag{87}$$

and

(ii) *Para* molecules:

$$\mathcal{Z}(\text{nuclear})\mathcal{Z}(\text{rotation}) = i(2i+1)\sum_{J\text{even}}(2J+1)e^{-J(J+1)\theta/T}, \tag{88}$$

where $\mathcal{Z}$ (nuclear) is the partition function for nuclear spin and θ is the rotational temperature. It should be noted from Eq. (87) that the energy of the ground state ($J = 1$ for *ortho* molecules) is not, in this case, equal to zero.

As an example, consider now the hydrogen molecule. With $i = \frac{1}{2}$ for the protons, the statistical weights are given by $(i+1)(2i+1) = 3$ and $i(2i+1) = 1$ for the *ortho* and *para* species, respectively. As the translational and vibrational partition functions are identical for the two types of hydrogen, the partial-pressure ratio approaches three at high temperatures ($T \gg \theta$), as the two sums in Eqs. (87) and (89) are equal in this limit. At low-temperatures, as $T \to 0$K, the sum in Eq. (87) becomes equal to zero and at equilibrium only *para*-H_2 can exist. This, result, as expressed by the percentages of the two types of molecular hydrogen, are presented as a function of temperature in Fig. 3. However, the equilibrium between these two species is only very slowly established in the absence of a catalyst.

For molecular hydrogen the evaluation of the various thermodynamic quantities from the partition functions is direct. However, for the rotational contribution the partition function must be considered to be the weighted sum of those for the *ortho* and *para* species. For example, the rotational part of the heat capacity is calculated from Eq. (46) with $\tilde{C}_v(\text{rotation}) = \frac{1}{4}\tilde{C}_v(para) + \frac{3}{4}\tilde{C}_v(ortho)$. At ordinary temperatures the equilibrium constant for the reaction *ortho*-$H_2 \leftrightarrow$ *para*-H_2 has almost reached its limiting value of $\frac{1}{3}$, as determined by the quantity $i/(i+1)$. At lower temperatures the sums involved in the rotational partition functions must be evaluated term-by-term. The calculated

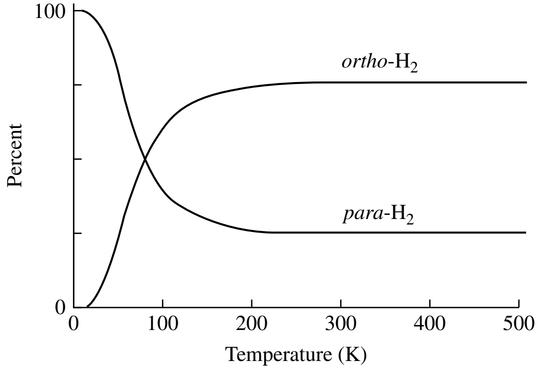


Fig. 3 Relative equilibrium percentages of *ortho*- and *para*-hydrogen as a function of temperature.

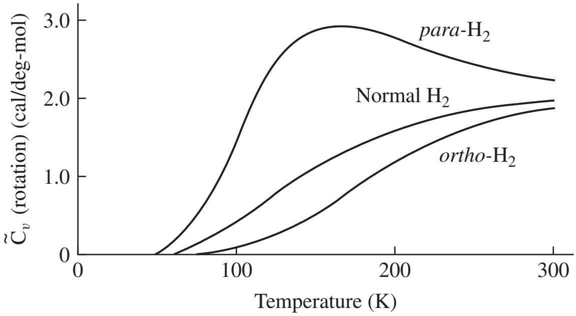


Fig. 4 Rotational contribution to the molar heat capacity $\tilde{C}_v$ for *ortho, para* and normal hydrogen. Note that 1cal/deg-mol = 4.18 J $K^{-1}mol^{-1}$.

rotational molar heat capacity is shown as a function of temperature in Fig. 4 for both *ortho*-H_2 and *para*-H_2, as well as for the equilibrium mixture.

If only the spins of the nuclei are altered in a given transition, the translational and vibrational contributions to the partition function are identical. Thus, for the reaction *ortho*-H_2 → *para*-H_2, the partial-pressure ratio at equilibrium is given by

$$K = \frac{[\textit{para}\text{-}H_2]}{[\textit{ortho}\text{-}H_2]} = \frac{\mathcal{Z}_{para}(\text{rotation})}{3\mathcal{Z}_{ortho}(\text{rotation})}. \tag{89}$$

However, it should be emphasized that the equilibrium between these two types of hydrogen is not easily established. In fact, under ordinary conditions, equilibrium is achieved only after a period of some years. It was for this reason that the early experimental determinations of the heat capacity of ordinary hydrogen were not in agreement with the values predicted from the elementary theory, in which the symmetry considerations presented above were not taken into account. Furthermore, the spectra of molecular hydrogen obtained at that time displayed anomalous line intensities that suggested the presence of two types of hydrogen, as indicated in this simple theoretical analysis.

It was subsequently shown that the presence of a catalyst accelerates the attainment of the *ortho–para* equilibrium. Measurements of properties such as the heat capacity and the thermal conductivity as a function of temperature then indicate that an equilibrium between the two species has been established.

PROBLEMS

1. Develop the series given by Eq. (18).
2. Evaluate $\int_1^n ln^m \, dx$ by parts and show that for $m = 1$ it becomes equal to $n \, ln \, n - n$ in the limit as $n \gg 1$.
3. Derive Eq. (28).
4. Verify Eq. (51).
5. Carry out the integration over J to obtain Eq. (59).
6. Derive the sum given in Eq. (65).
7. Verify Eq. (68).
8. Show that the function $f(x) = x^2 e^x/(e^x - 1)^2$ approaches zero as $x \to \infty$ and one as $x \to 0$. (Hint, use L'Hospital's rule.)
9. Verify Eq. (86).

11 Integral Transforms

Many pairs of functions $F(k)$ and $f(x)$ can be related by expressions of the form

$$F(k) = \int_{-\infty}^{\infty} f(x) K(k, x)\, \mathrm{d}x, \tag{1}$$

where the function $K(k, x)$ is known as the kernel. The function $F(k)$ is called the integral transform of the function $f(x)$ by the kernel $K(k, x)$. The operation described by Eq. (1) is sometimes referred to as the mapping of the function $f(x)$ in x space into another function $F(k)$ in k space. It is important to note that the variables x and k have reciprocal dimensions. Thus, for example if x has dimensions of frequency, k has dimensions of time in this case. Similarly, if x is a distance, say, in a crystal, k is a "distance" in the reciprocal lattice (see Section 4.6).

11.1 THE FOURIER TRANSFORM

By far the most useful integral transform in chemistry and physics is that of Fourier, *viz.*

$$F(k) = \int_{-\infty}^{\infty} f(x) e^{2\pi ikx}\, \mathrm{d}x, \tag{2}$$

where $i \equiv \sqrt{-1}$. Here, the kernel is complex, as given by Eulers's relation [Eq. (1-32)]. As any function of a real variable can be expressed as the sum of even and odd functions, *viz.*

$$f(x) = f_{\text{even}}(x) + f_{\text{odd}}(x), \tag{3}$$

where $f_{\text{even}}(x) = f_{\text{even}}(-x)$ and $f_{\text{odd}}(x) = -f_{\text{odd}}(-x)$, Eq. (2) becomes

$$F(k) = \int_{-\infty}^{\infty} f_{\text{even}}(x)\, cos(2\pi kx)\, \mathrm{d}x + i \int_{-\infty}^{\infty} f_{\text{odd}}(x)\, sin(2\pi kx)\, \mathrm{d}x \tag{4}$$

and $F(k)$ is complex. Clearly, as the functions sine and cosine are used to describe wave phenomena, Eq. (4) is employed in their analyses.

Other integral transforms are obtained with the use of the kernels e^{-kx} or x^{k-1}, among the infinite number of possibilities. The former yields the Laplace transform, which is of particular importance in the analysis of electrical circuits and the solution of certain differential equations. The latter was already introduced in the discussion of the gamma function (Section 5.5.4).

It is assumed that a given Fourier-transform operation, represented by

$$F(k) = \mathscr{F} f(x), \tag{5}$$

possesses an inverse such that

$$f(x) = \mathscr{F}^{-1} F(k). \tag{6}$$

Two functions that are related by Eqs. (5) and (6) are known as a transform pair. Thus, for example, the inverse of Eq. (2) is given by

$$f(x) = \int_{-\infty}^{\infty} F(k) e^{-2\pi i k x} \, \mathrm{d}k, \tag{7}$$

which is also a Fourier transform. It should be noted that a Fourier transform and its inverse, as defined here, are symmetrical – aside from the differing signs in the exponents.*

11.1.1 Convolution

The Fourier transform of the product of two functions is given by

$$\mathscr{G}(k) = \int_{-\infty}^{\infty} f(x) g(x) e^{2\pi i k x} \, \mathrm{d}x. \tag{8}$$

The Fourier transform of $f(x)$ can be written as

$$F(h) = \int_{-\infty}^{\infty} f(x) e^{2\pi i h x} \, \mathrm{d}x, \tag{9}$$

and its inverse as

$$f(x) = \int_{-\infty}^{\infty} F(h) e^{-2\pi i h x} \, \mathrm{d}h. \tag{10}$$

The substitution of Eq. (10) in Eq. (8) yields the relation

$$\mathscr{G}(k) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(h) g(x) e^{2\pi i (k-h) x} \, \mathrm{d}h \, \mathrm{d}x. \tag{11}$$

---

*If the factor 2π is not included in the exponents, its inverse will appear as a factor before either integral.

By reversing the order of integration in Eq. (11)

$$\mathcal{C}(k) = \int_{-\infty}^{\infty} F(h) \int_{-\infty}^{\infty} g(x) e^{2\pi i (k-h) x} \, \mathrm{d}x \, \mathrm{d}h \tag{12}$$

and it is apparent that the inner integral is just the Fourier transform $\mathscr{F} g(x) = G(k-h)$. Thus,

$$\mathcal{C}(k) = \int_{-\infty}^{\infty} F(h) G(k-h) \, \mathrm{d}h \equiv F(k) \bigstar G(k). \tag{13}$$

The star in Eq. (13) specifies the convolution operation. This operation arises in many branches of physics, chemistry and engineering. It is sometimes referred to as "scanning" or "smoothing". In general the convolution of one function with another is carried out by accumulating the result of successively displacing one function with respect to the other along the abscissa. Examples of the application of this principle are described later in this chapter. Numerical methods of evaluating the convolution operation are summarized in Chapter 13.

It is appropriate to note here that convolution is commutative, *viz.*

$$F(k) \bigstar G(k) = G(k) \bigstar F(k) \tag{14}$$

(problem 1). Furthermore, it is associative, as

$$F(k) \bigstar [G(k) \bigstar H(k)] = [F(k) \bigstar G(k)] \bigstar H(k) \tag{15}$$

and distributive under addition, with

$$F(k) \bigstar [G(k) + H(k)] = F(k) \bigstar G(k) + F(k) \bigstar H(k). \tag{16}$$

A particular case of convolution is that of a function with itself. From Eq. (13) this self-convolution can be expressed by

$$F(k) \bigstar F(k) = \int_{-\infty}^{\infty} F(h) F(k-h) \, \mathrm{d}h. \tag{17}$$

It can be considered to represent the cumulative effect of scanning the function $F(k)$ over itself.

11.1.2 Fourier transform pairs

In the following sections the most important, and relatively simple, transform pairs will be described. They have been chosen, as they represent those that are routinely applied in physical chemistry. Specifically, they are the functions that form the basis of the Fourier-transform techniques that are currently employed in virtually all areas of atomic and molecular spectroscopy. These functions are all even, or can be made so, thus their Fourier transforms are real.

The function "boxcar"

Consider the rectangular function shown in Fig. 1. It is a function that is equal to zero outside the region defined by the limits $-\ell$ and $+\ell$. However, within this region it has a constant value, determined by the condition that the integral over the function is equal to unity. That is to say, the function has been normalized. The rectangular function $\sqcap(x/2\ell)$ is often referred to as the "boxcar".

The Fourier transform of the normalized boxcar function can be obtained as follows. With $f(x) = (1/2\ell)\ \sqcap(x/2\ell)$, Eq. (2) can be written in the form

$$\mathscr{F}[(1/2\ell)\ \sqcap(x/2\ell)] = \frac{1}{2\ell}\int_{-\infty}^{\infty} \sqcap(x/2\ell)\, cos(2\pi kx)\, \mathrm{d}x \tag{18}$$

$$= \frac{1}{\ell}\int_{0}^{\ell} cos(2\pi kx)\, \mathrm{d}x, \tag{19}$$

as the integrand in Eq. (18) is an even function of x. The result given by Eq. (19) leads to

$$\mathscr{F}(1/2\ell)\ \sqcap(x/2\ell) = \frac{sin(2\pi k\ell)}{2\pi k\ell} \equiv sinc(2\pi k\ell) \tag{20}$$

(problem 2). This Fourier transform pair is illustrated in Fig. 1.

The sharp cutoff at the limits $-\ell$ and ℓ, as illustrated by the boxcar function, often occurs in the frequency domain. In this case the boxcar acts as a low-pass filter in applications in electronics. All frequencies below $|\ell|$ are unaltered, while in this ideal case all higher ones are suppressed.

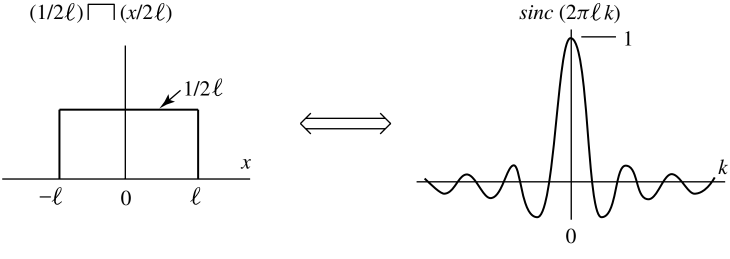


Fig. 1 The function boxcar and its Fourier transform, *sinc*.

The function triangle

This function can be defined by

$$\wedge(x/\ell) = \begin{cases} 0, & |x| > \ell \\ 1 - |x/\ell|, & |x| < \ell \end{cases} . \tag{21}$$

It is plotted with the appropriate normalizing factor in Fig. 2. The Fourier transform of this function can be evaluated as

$$\frac{1}{\ell}\int_{-\infty}^{\infty} \wedge(x/\ell)\, cos(2\pi kx)\, \mathrm{d}x = \frac{2}{\ell}\int_0^{\ell} \left(1 - \frac{x}{\ell}\right) cos(2\pi kx)\, \mathrm{d}x \tag{22}$$

$$= sinc^2(\pi k\ell) \tag{23}$$

(problem 3).

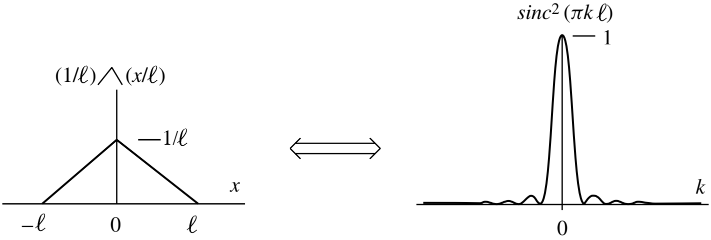


Fig. 2 The triangle function and its Fourier transform, *sinc*^{2}.

The Fourier transform of the triangle, as given by Eq. (23), is then

$$(1/\ell)\mathscr{F} \wedge (x/\ell) = sinc^2(\pi k\ell), \tag{24}$$

as shown in Fig. 2. From this result it is apparent that the triangle function is the result of the self-convolution of two boxcars. A well-known example in optics is provided by a monochromator in which the image of the rectangular entrance slit is scanned over a rectangular exit slit. The resulting triangle is referred to as the slit function (see problem 4).

Gauss's function

The Gaussian function was discussed in Section 3.4.5. When normalized it takes the form

$$f(x) = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha x^2}. \tag{25}$$

The Fourier transform of the Gaussian is given by

$$F(k) = \sqrt{\frac{\alpha}{\pi}} \int_{-\infty}^{\infty} e^{-\alpha x^2} \, cos(2\pi kx) \, \mathrm{d}x \tag{26}$$

$$= e^{-\pi^2 k^2/\alpha}. \tag{27}$$

Thus, the function of Gauss is its own Fourier transform, as shown in Fig. 3 (see problem 5).

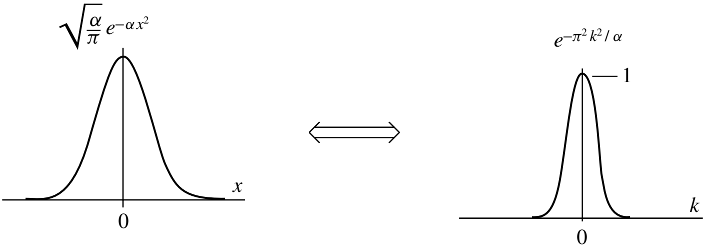


Fig. 3 The normalized Gaussian and its Fourier transform, a Gaussian in k space.

Exponential decay: The Lorentz profile[*]

A phenomenon that exhibits an exponential decay – for example, in time – results in a Lorentian function in the frequency domain. The Fourier transform of the normalized function

$$f(x) = \frac{\beta}{2} e^{-\beta|x|} \tag{28}$$

is given by

$$F(k) = \frac{\beta}{2} \int_{-\infty}^{\infty} e^{-\beta|x|} \, cos(2\pi kx) \, \mathrm{d}x \tag{29}$$

$$= \frac{1}{1 + (4\pi^2 k^2/\beta^2)} \tag{30}$$

(see problem 6). This function, is sometimes referred to as the function of Cauchy. It is, along with the Gaussian [Eq. (25)], often used to describe the profile of an observed spectroscopic feature, *e.g.* the “bandshape”.

There is a fundamental interest in the profiles of spectral bands. As they are functions of frequency, it should be clear from the arguments presented

---

[*] Hendrik Antoon Lorentz, Dutch physicist (1853–1928).

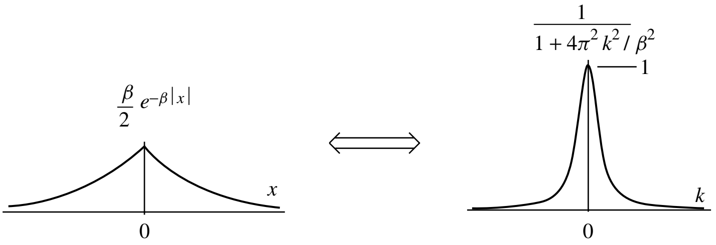


Fig. 4 The normalized exponential and its Fourier transform, the Lorentzian.

above that their Fourier transforms are functions of time. Thus the analyses of observed bandshapes provides molecular dynamic information, that is, quantitative descriptions of the time evolution of molecular interactions.

The delta function of Dirac and the "Shah"

The Dirac delta function represents an intense impulse of very short time duration. An example is the "hit" of a baseball by the bat. From a mathematical point of view this function can be defined by the relations

$$\delta(x) = 0, \quad \text{if } x \neq 0 \tag{31}$$

and

$$\int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x = 1. \tag{32}$$

Dirac's $\delta(x)$ cannot be considered to be a function in the usual sense. Although in principle it can only be written under an integral sign, Eq. (32) can be interpreted as a limit, *viz.* $\lim_{t\to 0} \int_{-\infty}^{\infty} 1/2\ell \;\sqcap(x/2\ell)\,\mathrm{d}x$. The integrand in this expression has unit area. Thus, as $\ell \to 0$ the function is limited to the region near the origin and, as it becomes narrower, its height increases to compensate. This property of the delta function is illustrated in Fig. 5.

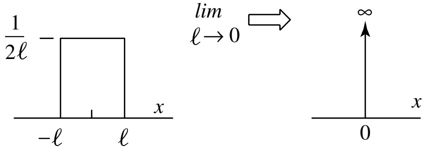


Fig. 5 The delta function as the limit of the boxcar.

The Fourier transform of the function $\delta(x - \ell)$ is given by

$$F(k) = \int_{-\infty}^{\infty} \delta(x - \ell)\, cos(2\pi k\ell)\, \mathrm{d}x. \tag{33}$$

Therefore, with $\ell = 0$ it is apparent that the Fourier transform of the delta function is equal to unity.

Two properties of the delta function are of particular interest:

(i) If it is multiplied by a function $f(x)$, its effect is to yield only the value of the function at the point where the delta function is nonzero. Thus,

$$f(x)\delta(x - \ell) = f(\ell) \tag{34}$$

and the function $f(x)$ has been "sampled" at the point $x = \ell$.

(ii) If it is convoluted with the function $f(x)$, it acts as a shifting operator. Then,

$$f(x)\bigstar\delta(x - \ell) = \int_{-\infty}^{\infty} f(x')\delta(x - x' - \ell)\, \mathrm{d}x' \tag{35}$$

$$= f(x - \ell) \tag{36}$$

and the function $f(x)$ has been shifted to $f(x - \ell)$.

The principle of sampling suggested in the previous paragraph can be generalized in the form

$$(1/\ell)\ \text{Ш}\ (x/\ell) = \sum_{n=-\infty}^{\infty} \delta(x - n\ell), \tag{37}$$

where n is an integer. The function $\text{Ш}\,(x/\ell)$ (Hebrew: *shah*) has been introduced to represent the "sampling comb". Multiplication of a function $f(x)$ by $\text{Ш}\,(x/\ell)$ selects its values at equal intervals, as the spacing ℓ between successive "teeth" in the comb is constant. Thus,

$$(1/\ell)\ \text{Ш}\ (x/\ell) \times f(x) = \sum_{n=-\infty}^{\infty} f(n\ell). \tag{38}$$

The values of $f(x)$ are only retained at each particular point where $x = n\ell$. Furthermore, the convolution of a function $f(x)$ with the sampling comb results in a function

$$(1/\ell)\ \text{Ш}\ (x/\ell)\bigstar f(x) = \sum_{n=-\infty}^{\infty} f(x - n\ell), \tag{39}$$

which represents the endless replication of the original function at equal intervals. This relation is the origin of what is known in Fourier-transform spectroscopy as "aliasing".

Like the Gaussian, discussed above, the function shah is its own Fourier transform. Thus,

$$F(k) = \frac{1}{\ell}\int_{-\infty}^{\infty} \text{Ш}\,(x/\ell)e^{2\pi ikx}\,\mathrm{d}x \tag{40}$$

$$= \frac{1}{\ell}\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x-n\ell)e^{2\pi ikx}\,\mathrm{d}x \tag{41}$$

$$= \frac{1}{\ell}\sum_{n=-\infty}^{\infty}\delta\left(k-\frac{n}{\ell}\right) = \text{Ш}\,(k\ell). \tag{42}$$

This relation is illustrated in Fig. 6.

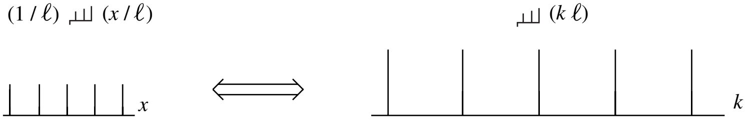


Fig. 6 The relation $\mathscr{F}(1/\ell)\ \text{Ш}\ (x/\ell) = \text{Ш}\ (k\ell)$.

11.2 THE LAPLACE TRANSFORM

The Laplace transform can be defined by

$$F(s) = \mathscr{L}f(t) = \int_{0}^{\infty} e^{-st}f(t)\,\mathrm{d}t. \tag{43}$$

As such, with integration limits from zero to infinity, it is referred to as a "one-sided" transform. For simplicity, it will be assumed here that the variable s is real and positive. Again, as in the case of the Fourier transform, the variables s and t have reciprocal dimensions and the operator $\mathscr{L}$ is linear (see Section 7.1).

11.2.1 Examples of simple Laplace transforms

(i) If $f(t) = 1$,

$$\mathscr{L}(1) = \int_{0}^{\infty} e^{-st}\,\mathrm{d}t = \frac{1}{s}\,. \quad t > 0 \tag{44}$$

(ii) If $f(t) = e^{kt}, \quad t > 0$ (45)

$$\mathscr{L}(e^{kt}) = \int_0^\infty e^{-st} e^{kt}\, \mathrm{d}t = \frac{1}{s-k} \ . \quad s > k \tag{46}$$

(iii) From (ii) the relations

$$\mathscr{L}(cosh\ kt) = \frac{s}{s^2 - k^2} \quad s > k \tag{47}$$

and

$$\mathscr{L}(sinh\ kt) = \frac{k}{s^2 - k^2} \quad s > k \tag{48}$$

can be easily derived [problem 7; see Eqs. (1-44) and (1-45)].

(iv) With the aid of Eqs. (1-46) and (1-47) the corresponding relations for the circular functions can be found, *viz.*

$$\mathscr{L}(cos\, kt) = \frac{s}{s^2 + k^2} \quad s > k \tag{49}$$

and

$$\mathscr{L}(sin\, kt) = \frac{k}{s^2 + k^2} \quad s > k \tag{50}$$

(problem 8).

$$\text{(v)}\, \mathscr{L}(t^n) = \int_0^\infty e^{-st} t^n\, \mathrm{d}t \tag{51}$$

$$= \frac{n!}{s^{n+1}} = \frac{\Gamma(n+1)}{s^{n+1}}, \quad s > 0,\ n > -1 \tag{52}$$

where n is an integer (see Section 5.5.4 for the definition of the gamma function).

(vi) If s is replaced by $(s - a)$ in Eq. (43), the Laplace transform becomes

$$F(s-a) = \int_0^\infty e^{-(s-a)t} f(t)\, \mathrm{d}t \tag{53}$$

$$= \mathscr{L}[e^{at} f(t)]. \tag{54}$$

This relation can be employed to obtain other Laplace transforms. For example, with the use of Eqs. (49) and (50),

$$\mathscr{L}(e^{at}\,cos\,kt) = \frac{s-a}{(s-a)^2+k^2} \tag{55}$$

and

$$\mathscr{L}(e^{at}\,sin\,kt) = \frac{k}{(s-a)^2+k^2}\ . \tag{56}$$

Many other Laplace transforms can be derived in this way. Extensive tables of Laplace transforms are available and are of routine use, particularly by electronics engineers.

11.2.2 The transform of derivatives

The Laplace transform of the derivative of a function $f(t)$ is given by

$$\mathscr{L}\left[\frac{\mathrm{d}f(t)}{\mathrm{d}t}\right] = \int_0^\infty e^{-st}\frac{\mathrm{d}f(t)}{\mathrm{d}t}\,\mathrm{d}t. \tag{57}$$

Integration by parts yields the result

$$\mathscr{L}\left[\frac{\mathrm{d}f(t)}{\mathrm{d}t}\right] = e^{-st}f(t)\Big|_0^\infty + s\int_0^\infty e^{-st}f(t)\,\mathrm{d}t \tag{58}$$

$$= -f(0) + s\mathscr{L}f(t). \tag{59}$$

The transforms of higher derivatives can be found by the same method, *e.g.*

$$\mathscr{L}\left[\frac{\mathrm{d}^2 f(t)}{\mathrm{d}t^2}\right] = s^2\mathscr{L}f(t) - sf(0) - \frac{\mathrm{d}f}{\mathrm{d}t}(0) \tag{60}$$

(problem 10).

A simple example of the application of Eq. (60) is provided by the function *sin kt*. As

$$\frac{\mathrm{d}^2\,sin\,kt}{\mathrm{d}t^2} = -k^2\,sin\,kt \tag{61}$$

and

$$\mathscr{L}\left[\frac{\mathrm{d}^2\,sin\,kt}{\mathrm{d}t^2}\right] = -k^2\mathscr{L}\,sin\,kt \tag{62}$$

$$= s^2\mathscr{L}\,sin\,kt - sin(0) - \left.\frac{\mathrm{d}\,sin\,kt}{\mathrm{d}t}\right|_{t=0}. \tag{63}$$

Then, as $(\mathrm{d}\,sin\,kt/\mathrm{d}t)|_{t=0} = k$ and $sin(0) = 0$,

$$\mathscr{L}\,sin\,kt = \frac{k}{s^2 + k^2}\,, \tag{64}$$

in agreement with Eq. (50).

11.2.3 Solution of differential equations

It should be evident that the expressions for the Laplace transforms of derivatives of functions can facilitate the solution of differential equations. A trivial example is that of the classical harmonic oscillator. Its equation of motion is given by Eq. (5-33), namely,

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \frac{\kappa}{m}x = 0. \tag{65}$$

By taking the Laplace transform, this equation becomes

$$\mathscr{L}\left[\frac{\mathrm{d}^2x}{\mathrm{d}t^2}\right] + \frac{\kappa}{m}\mathscr{L}(x) = 0. \tag{66}$$

Note that the use of Eq. (60) for the transform of the second derivative includes the initial conditions on the solution to the problem. The introduction of the initial conditions at this point is to be compared with the procedure employed earlier (see Section 5.2.2). If the conditions $x(0) = x_0$ and $\mathrm{d}x/\mathrm{d}t = 0$ are applied,

$$s^2X(s) - sx_0 + \frac{\kappa}{m}X(s) = 0 \tag{67}$$

and

$$X(s) = \frac{s}{s^2 + (\kappa/m)^2}x_0. \tag{68}$$

Thus, from Eq. (49) the solution is given by

$$x(t) = x_0\,cos\,\omega_0 t \tag{69}$$

where $\omega_0^2 = \kappa/m$.

As a somewhat more complicated example, consider the electrical circuit of the damped oscillator shown in Fig. 5-3. The charge $q(t)$ is determined by Eq. (5-45), namely,

$$\frac{\mathrm{d}^2q(t)}{\mathrm{d}t^2} + \frac{R}{L}\frac{\mathrm{d}q(t)}{\mathrm{d}t} + \frac{1}{LC}q(t) = 0, \tag{70}$$

where R is the resistance, L the inductance and C the capacitance. The capacitance will be assumed to have an initial charge q_0, at which time there is no current flowing in the circuit. Thus, the initial conditions are $q(0) = q_0$ and $(\mathrm{d}q/\mathrm{d}t)(0) = 0$. The Laplace transform then yields the relation

$$s^2(X)(s) - sq_0 + \frac{R}{L}[sX(s) - q_0] + \frac{1}{LC}X(s) = 0. \tag{71}$$

Then,

$$X(s) = q_0 \frac{s + R/L}{s^2 + s(R/L) + 1/LC} \tag{72}$$

and the condition for oscillation is

$$\omega_1^2 = \frac{1}{LC} - \frac{R^2}{4L^2} > 0. \tag{73}$$

This result is equivalent to that derived in Section 5.2.3 for the mechanical analog (problem 11). Equation (72) can be written

$$X(s) = q_0 \frac{s + R/L}{(s + R/2L)^2 + \omega_1^2} = q_0 \frac{s + R/2L}{(s + R/2L)^2 + \omega_1^2} + q_0 \frac{R/2L}{(s + R/2L)^2 + \omega_1^2} \tag{74}$$

and, with the aid of Eqs. (49) and (50), the solution becomes

$$q(t) = q_0 e^{-Rt/2L} \left(cos\, \omega_1 t + \frac{R}{2L\omega_1} sin\, \omega_1 t \right), \tag{75}$$

in agreement with Eq. (5-43). Here again, the initial conditions are specified at the outset (problem 12).

11.2.4 Laplace transforms: Convolution and inversion

The convolution and general properties of the Fourier transform, as presented in Section 11.1, are equally applicable to the Laplace transform. Thus,

$$\int_{-\infty}^{\infty} F(h)G(k - h)\,\mathrm{d}h \equiv F(k)\bigstar G(k), \tag{76}$$

where $F(k)$ and $G(k)$ are the Laplace transforms of $f(x)$ and $g(x)$, respectively.

The inversion of the Laplace transform presents a more difficult problem. From a fundamental point of view the inverse of a given Laplace transform is known as the Bromwich integral.* Its evaluation is carried out by application

---

*Thomas John I'anson Bromwich, English mathematician (1875–1929).

of the theory of residues. As contour integration is not treated in this book, the reader is referred to more advanced texts for the explanation of this method.

In practice, the inverse Laplace transformations are obtained by reference to the rather extensive tables that are available. It is sometimes useful to develop the function in question in partial fractions, as employed in Section 3.3.3. The resulting sum of integrals can often be evaluated with the use of the tables.

In principle, numerical methods can be employed to evaluate inverse Laplace transforms. However, the procedure is subject to errors that are often very large–even catastrophic.

11.2.5 Green's functions*

The introduction of these somewhat mysterious functions allows certain differential equations to be converted into equivalent integral equations. Although the method is particularly useful in its application to partial differential equations, it will be illustrated here with the aid of a relatively simple example, the forced vibrations of a classical oscillator.

Consider first the inhomogeneous differential equation as given by Eq. (5-57). For simplicity, assume here that the oscillator is not damped; hence, $h = 0$. The problems to be treated are now represented by the differential equation

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \omega_0^2 x = \phi(t). \tag{77}$$

The function $\phi(t)$, aside from a constant, expresses the time-dependent force acting on the harmonic oscillator and $\omega_0 = \sqrt{\kappa/m}$ is the angular frequency of the system (Section 5.3.3).

Now consider the external force acting on the system to be composed of a series of instantaneous impacts, each of which can be expressed mathematically by a delta function. The response of the system can then be represented by a function $G(t)$. The differential equation to be solved then takes on the form

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} G(t, t') + \omega_0^2 G(t, t') = \delta(t - t'). \tag{78}$$

The function $\delta(t - t')$ corresponds to an impact on the system at the instant $t = t'$. The function $G(t, t')$ is known as a Green's function. It has been implied here that the forcing function $\phi(t)$ can be represented by a sum of such delta functions, as given by

$$\phi(t) = \int_0^\infty \phi(t')\delta(t - t')\,\mathrm{d}t' \qquad t > t' > 0 \tag{79}$$

---

*George Green, English mathematician (1793–1841).

[See Eq. (35) and Eq. (36) with $x - \ell$ replaced by t]. In effect, the delta function samples the forcing function at each point in time. The proof of Eq. (79) constitutes problem 13.

The solution of Eq. (78) can be obtained with the use of the Laplace transform. However, it is first necessary to develop the expression for the Laplace transform of the delta function, as given on the right-hand side of Eq. (78). With the use of the definition of the Laplace transform [Eq. (43)] and $f(t) = \delta(t - t')$, the desired result becomes

$$\mathscr{L}\delta(t - t') = e^{-st'}. \tag{80}$$

The Laplace transform of Eq. (78) can then be written as

$$(s^2 + \omega_0^2)\mathscr{L}G(t, t') = \mathscr{L}\delta(t - t') = e^{-st}, \tag{81}$$

where Eq. (60) has been employed with the initial conditions $G(0, 0) = G'(0, 0) = 0$. Thus, with the use of Eqs. (50) and (54) to obtain the inverse Laplace transforms, the solution is given by

$$G(t, t') = \mathscr{L}^{-1}\left[\frac{e^{-st}}{s^2 + \omega_0^2}\right] = \frac{1}{\omega_0} sin[\omega_0(t - t')] \tag{82}$$

(see problem 14). This result is easily verified by substitution in Eq. (78). Once the Green's function has been found for this type of problem, the solution of Eq. (77) for a specific forcing function $\phi(t')$ can, at least in principle, be obtained by direct integration, namely,

$$x(t) = \int_0^\infty G(t, t')\phi(t')\,\mathrm{d}t' \tag{83}$$

(problem 15). Note that Eq. (83) is an integral equation with $G(t, t')$ the kernel.

As a simple example of the general method outlined above, consider the vibrations of the harmonic oscillator under the forcing function $\phi(t') = F_0\, sin\, \omega t'$, as in Eq. (5-58). Thus, Eq. (83) becomes

$$x(t) = \frac{F_0}{\omega_0}\int_0^\infty sin[\omega_0(t - t')]\, sin\, \omega t'\,\mathrm{d}t'. \tag{84}$$

This integral can be evaluated with the aid of the appropriate trigonometric relations. Furthermore, the upper limit can be replaced by t, as $t' < t$. The result,

$$x(t) = F_0\frac{sin\, \omega_0 t}{\omega^2 - \omega_0^2} \tag{85}$$

is (aside from the factor F_0) the same as Eq. (5-59) for the special case where $h = 0$. This solution to the inhomogeneous differential equation is referred to as the "steady-state" solution, as contrasted to the transient one [Eq. (5-34)] that becomes negligible with increasing time. The catastrophic behavior of the forced oscillator if ω is close to ω_0 was discussed in Section 5-33.

PROBLEMS

1. Show that $F(k)\bigstar G(k) = G(k)\bigstar F(k)$.
2. Derive the expression for the Fourier transform of the boxcar [Eq. (20)].
3. Verify Eq. (23).
4. To illustrate the self-convolution operation, draw two identical boxcars and evaluate the area in common as a function of their relative separation along the abscissa.
5. Show that the Fourier transform of a Gaussian is also a Gaussian.
6. Verify Eq. (30).
7. Verify Eqs. (44) to (48).
8. Verify Eqs. (49) and (50).
9. Derive Eqs. (52) to (56).
10. Derive Eq. (60).
11. Compare Eq. (73) with the corresponding result for the damped mechanical oscillator.
12. Verify Eq. (75).
13. Derive Eq. (79).
14. Substitute Eq. (82) to show that it is a solution to Eq. (78).
15. Demonstrate that Eq. (73) is a solution to Eq. (77).
16. Carry out the integration indicated in Eq. (84) to obtain Eq. (85).

12 Approximation Methods in Quantum Mechanics

12.1 THE BORN–OPPENHEIMER APPROXIMATION

In this section the Born–Oppenheimer approximation will be presented in what is necessarily a very simplified form. It has already been introduced without justification in Section 6.5. It is certainly the most important – and most satisfactory – approximation in quantum mechanics, although its rigorous derivation is far beyond the level of this book. Consider, therefore, the following argument.

A stationary state of a polyatomic molecule can be described in quantum mechanics by a wavefunction ψ and an energy ε. Thus, according to Schrödinger,

$$\hat{H}\psi = \varepsilon\psi, \tag{1}$$

where $\hat{H}$ is the total Hamiltonian of the system; it depends on the electronic, nuclear and spin coordinates. Thus, the Hamiltonian is a function of a large number of independent variables that must be separated, at least approximately, to obtain equations that can be solved for variables of each type.

It is usually assumed that the spins that are included in the Hamiltonian can be removed. The variables involving spins of the various particles will thus be eliminated, and the Hamiltonian remains then a function of the positions of the nuclei and the electrons.

The separation of the electronic and nuclear motions depends on the large difference between the mass of an electron and the mass of a nucleus. As the nuclei are much heavier, by a factor of at least 1800, they move much more slowly. Thus, to a good approximation the movement of the electrons in a polyatomic molecule can be assumed to take place in the environment of the nuclei that are fixed in a particular configuration. This argument is the physical basis of the Born–Oppenheimer approximation.

The Hamiltonian for a system composed of a number of nuclei and electrons can be written in the form

$$\hat{H} = -\sum_j \frac{\hbar^2}{2M_j}\nabla_j^2 - \sum_i \frac{\hbar^2}{2m_i}\nabla_i^2 + V_{jj'} + V_{ii'} + V_{ij}, \tag{2}$$

where

$$V_{jj} = \frac{1}{4\pi\varepsilon_0}\sum_{jj',\ j>j'} \frac{Z_j Z_{j'} e^2}{R_{jj'}}, \tag{3}$$

$$V_{ii'} = \frac{1}{4\pi\varepsilon_0}\sum_{ii',\ i>i'} \frac{e^2}{r_{ii'}} \tag{4}$$

and

$$V_{ij} = -\frac{1}{4\pi\varepsilon_0}\sum_{ij} \frac{Z_j e^2}{r_{ij}}. \tag{5}$$

The first two terms in Eq. (2) represent the kinetic energy of the nuclei and the electrons, respectively. The remaining three terms specify the potential energy as a function of the interaction between the particles. Equation (3) expresses the potential function for the interaction of each pair of nuclei. In general, this sum is composed of terms that are given by Coulomb's law for the repulsion between particles of like charge. Similarly, Eq. (4) corresponds to the electron–electron repulsion. Finally, Eq. (5) is the potential function for the attraction between a given electron (i) and a nucleus (j).

It is useful in the present context to separate the Hamiltonian of Eq. (2) into two parts, as given by

$$\hat{H} = \hat{H}_{elec} + \hat{H}_{nucl} = -\sum_i \frac{\hbar^2}{2m_i}\nabla_i^2 + V_{ii'} + V_{ij} - \sum_j \frac{\hbar^2}{2M_j}\nabla_j^2 + V_{jj'}. \tag{6}$$

It is then assumed that the wavefunction can be approximated by the relation $\psi = \psi_{elec}\psi_{nucl}$. The first three terms in Eq. (6) are referred to as the electronic part of the Hamiltonian, while the remaining two terms represent the nuclear Hamiltonian. The Schrödinger equation for the general problem can then be written as

$$\begin{aligned}\hat{H}(\psi_{elec}\psi_{nucl}) = -\sum_j \frac{\hbar^2}{2M_j}\nabla_j^2(\psi_{elec}\psi_{nucl}) - \sum_i \frac{\hbar^2}{2m_i}\nabla_i^2(\psi_{elec}\psi_{nucl})\\ + (V_{jj'} + V_{ii'} + V_{ij})\psi_{elec}\psi_{nucl} = \varepsilon\psi_{elec}\psi_{nucl}.\end{aligned} \tag{7}$$

As the first term on the right-hand side of Eq. (7) expresses the kinetic energy of the nuclei, it is assumed that the wavefunction for the electrons is not modified by the corresponding operator. In other words the movement of the nuclei takes place within the so-called electron cloud that is developed by the average positions of the ensemble of electrons. Inversely, in the second term of Eq. (7) the operator over the electronic coordinates has no effect on the nuclear wavefunction, as the nuclear positions are taken to be fixed over the period of time of electronic motion.

The imposition of the above approximations on Eq. (7), followed by division by ψ_{nucl}, leads to the relation

$$\frac{\psi_{elec}}{\psi_{nucl}}\sum_j\left(-\frac{\hbar^2}{2M_j}\nabla_j^2\psi_{nucl}\right)+\left[\sum_i\left(-\frac{\hbar^2}{2m_i}\nabla_i^2+V_{ii'}+V_{ij}\right)\psi_{elec}\right] - (\varepsilon - V_{jj'})\psi_{elec} = 0. \tag{8}$$

The quantity in square brackets can be readily identified as

$$\hat{H}_{elec}\psi_{elec} = \varepsilon_{elec}\psi_{elec}, \tag{9}$$

where $\varepsilon = \varepsilon_{elec} + \varepsilon_{nucl}$. With the aid of this substitution and multiplication by ψ_{nucl}/ψ_{elec}, the expression

$$\sum_j\left(-\frac{\hbar^2}{2M_j}\nabla_j^2\psi_{nucl}\right)+V_{jj'}\psi_{nucl}+(\varepsilon_{elec}-\varepsilon)\psi_{nucl}=0 \tag{10}$$

is obtained. If Eq. (10) is written in the form

$$(\hat{H}_{nucl}+\varepsilon_{elec})\psi_{nucl} = \varepsilon\psi_{nucl}, \tag{11}$$

it becomes evident that the electronic energy enters the Hamiltonian for the nuclear motion as an effective potential function. Thus, if the "electronic problem", as given by Eq. (9), has been solved as a function of the nuclear geometry, the resulting energy contributes to the potential function that governs internuclear displacements.

The argument presented in the preceding paragraph is perhaps easier to understand with reference to a diatomic molecule. In this case Eq. (9) is solved for each value of the (fixed) interatomic distance. The resulting electronic energy, as a function of this distance, is then substituted in Eq. (10) to yield the Schrödinger equation for the relative nuclear motion, as given by Eq. (6-70). The function $V_{jj'}$ is just the Coulombic repulsion between the two positively charged nuclei, while ε_{elec} is the potential function that describes the forces created by the electron cloud.

It should be emphasized that the Born–Oppenheimer approximation is an extremely good one. Only in certain questions in the interpretation of the molecular spectra of small molecules in the gas phase is it necessary to consider its inherent errors. Therefore, it will not be considered further. There are, however, other approximations that are important in many areas of physical chemistry and physics.

Several examples of the application of quantum mechanics to relatively simple problems have been presented in earlier chapters. In these cases it was possible to find solutions to the Schrödinger wave equation. Unfortunately, there are few others. In virtually all problems of interest in physics and chemistry, there is no hope of finding analytical solutions, so it is essential to develop approximate methods. The two most important of them are certainly perturbation theory and the variation method. The basic mathematics of these two approaches will be presented here, along with some simple applications.

12.2 PERTURBATION THEORY: STATIONARY STATES

12.2.1 Nondegenerate systems

In many problems for which no direct solution can be obtained, there is a wave equation which differs but slightly from one that can be solved analytically. As an example, consider the hydrogen atom, a problem that was resolved in Section 6.6. Suppose now that an electric field is applied to the atom. The energy levels of the atom are affected by the field, an example of the Stark effect.* If the field (due to the potential difference between two electrodes, for example) is gradually reduced, the system approaches that of the unperturbed hydrogen atom.

With the experiment described above in mind, represent the Hamiltonian of the unperturbed system by $\hat{H}^0$ and that of the perturbed system by

$$\hat{H} = \hat{H}^0 + \lambda \hat{H}'. \tag{12}$$

Assume that the perturbation, $\lambda \hat{H}'$ is small compared with $\hat{H}^0$, where λ is a parameter. As $\lambda \to 0$, the eigenvalues and eigenfunctions are those of the unperturbed system, as given by

$$\hat{H}^0 \psi_n^0 = \varepsilon_n^0 \psi_n^0 \,. \tag{13}$$

They are assumed to be known. Furthermore, it should be recalled that the eigenfunctions ψ_n^0 form a complete orthogonal set. The equation of interest is

$$(\hat{H}^0 + \lambda \hat{H}')\psi_n = \varepsilon_n \psi_n. \tag{14}$$

As ψ_n and ε_n are both functions of λ, they can be expanded in power series, *viz.*

---

*Johannes Stark, German physicist (1874–1957).

$$\psi_n = \psi_n^0 + \lambda\psi_n' + \lambda^2\psi_n'' + \ldots \tag{15}$$

and

$$\varepsilon_n = \varepsilon_n^0 + \lambda\varepsilon_n' + \lambda^2\varepsilon_n'' + \ldots\ . \tag{16}$$

It will be assumed here, with reasonable assurance, that these two series converge. The substitution of Eqs. (15) and (16) in Eq. (13) leads to the relation

$$\begin{aligned}\hat{H}^0\psi_n^0 &+ \lambda(\hat{H}'\psi_n^0 + \hat{H}^0\psi_n') + \lambda^2(\hat{H}'\psi_n' + \hat{H}^0\psi_n'') + \cdots \\ &= \varepsilon_n^0\psi_n^0 + \lambda(\varepsilon_n'\psi_n^0 + \varepsilon_n^0\psi_n') + \lambda^2(\varepsilon_n''\psi_n^0 + \varepsilon_n'\psi_n' + \varepsilon_n^0\psi_n'') + \ldots\ .\end{aligned} \tag{17}$$

The coefficients of the various powers of λ can now be collected to yield a series of equations, namely,

$$\lambda^0 : \hat{H}^0\psi_n^0 = \varepsilon_n^0\psi_n^0 \tag{18}$$

$$\lambda^1 : (\hat{H}^0 - \varepsilon_n^0)\psi_n' = \varepsilon_n'\psi_n^0 - \hat{H}'\psi_n^0 \tag{19}$$

$$\lambda^2 : (\hat{H}^0 - \varepsilon_n^0)\psi_n'' = \varepsilon_n''\psi_n^0 + \varepsilon_n'\psi_n' - \hat{H}'\psi_n' \tag{20}$$

$$\vdots\ .$$

It has been assumed above that Eq. (18) has been solved. In principle, the resulting eigenvalues and eigenfunctions can then be substituted in Eq. (19) to yield the first-order corrections, and so on, for higher orders of approximation.

12.2.2 First-order approximation

To resolve Eq. (19) an expansion of ψ_n' is made in terms of the zero-order eigenfunctions. Thus,

$$\psi_n' = \sum_\ell a_\ell\psi_\ell^0 \tag{21}$$

and

$$\hat{H}^0\psi_n' = \hat{H}^0\sum_\ell a_\ell\psi_\ell^0 = \sum_\ell a_\ell\varepsilon_\ell^0\psi_\ell^0\ . \tag{22}$$

Equation (19) is then written as

$$\sum_\ell a_\ell(\varepsilon_\ell^0 - \varepsilon_n^0)\psi_\ell^0 = (\varepsilon_n' - \hat{H}')\psi_n^0\ . \tag{23}$$

The orthonormal properties of the functions ψ_ℓ^0 can now be employed to obtain the desired result. Equation (23) is multiplied by $\psi_n^{0\star}$ and integrated over all space. The left-hand side of Eq. (23) becomes

$$\int \psi_n^{0\star} \sum_\ell a_\ell(\varepsilon_\ell^0 - \varepsilon_n^0)\psi_\ell^0 \, \mathrm{d}\tau = \sum_\ell a_\ell(\varepsilon_\ell^0 - \varepsilon_n^0) \int \psi_n^{0\star} \psi_\ell^0 \, \mathrm{d}\tau = 0. \tag{24}$$

The final result is obtained because if $\ell = n$, $\varepsilon_\ell^0 - \varepsilon_n^0 = 0$. Or, if $\ell \neq n$, $\int \psi_n^{0\star} \psi_\ell^0 \, \mathrm{d}\tau = 0$, as the functions are orthogonal. After the same operations the right-hand side of Eq. (23) is given by

$$\int \psi_n^{0\star} (\varepsilon_n' - \hat{H}')\psi_n^0 \, \mathrm{d}\tau = \varepsilon_n' \int \psi_n^{0\star} \psi_\ell^0 \, \mathrm{d}\tau - \int \psi_n^{0\star} \hat{H}'\psi_\ell^0 \, \mathrm{d}\tau \tag{25}$$

and, as the zero-order wavefunctions are normalized, Eq. (25) is simply

$$\varepsilon_n' = \int \psi_n^{0\star} \hat{H}'\psi_n^0 \, \mathrm{d}\tau. \tag{26}$$

This result stated in words is that ε_n', the first-order correction to energy of the system in a given state n, is just the average value of the perturbation of the Hamiltonian. Equation (26) can be written in the more compact notation of Dirac as

$$\varepsilon_n' = \left\langle n \left| \hat{H}' \right| n \right\rangle. \tag{27}$$

It is often of interest to calculate the corresponding first-order correction to the wavefunctions. The necessary expression can be obtained by returning to Eq. (23). If this equation is now multiplied on each side by $\psi_j^{0\star}$ and the result integrated over all space, the left side will vanish, as before, unless $j = \ell \neq n$. With the application of these conditions, the result is

$$a_j(\varepsilon_j^0 - \varepsilon_n^0) = \varepsilon_n' \int \psi_j^{0\star} \psi_n^0 \, \mathrm{d}\tau - \int \psi_j^{0\star} \hat{H}'\psi_n^0 \, \mathrm{d}\tau. \tag{28}$$

As $j \neq n$, the first term on the right-hand side of Eq. (28) is equal to zero and the coefficients in Eq. (21) are given by

$$a_j = -\frac{\int \psi_j^{0\star} \hat{H}'\psi_n^0 \, \mathrm{d}\tau}{\varepsilon_j^0 - \varepsilon_n^0} = -\frac{\left\langle j \left| \hat{H}' \right| n \right\rangle}{\varepsilon_j^0 - \varepsilon_n^0}. \tag{29}$$

The first-order wavefunctions are then

$$\psi_n = \psi_n^0 + \psi_n' = \psi_n^0 - \sum_{j=0}^{\infty}{}' \frac{\left\langle j \left| \hat{H}' \right| n \right\rangle}{\varepsilon_j^0 - \varepsilon_n^0} \psi_j^0 . \tag{30}$$

The prime on the summation in Eq. (30) indicates that the subscript $j = n$ is excluded. It should be noted that the first-order energy given by Eq. (27)

depends only on the diagonal elements of the matrix of $\hat{H}'$. However, the first-order wavefunctions given by Eq. (30) are determined by its off-diagonal elements.

12.2.3 Second-order approximation

Returning to Eq. (20), the coefficients of λ^2 yield the relation

$$(\hat{H}^0 - \varepsilon_n^0)\psi_n'' = (\varepsilon_n' - \hat{H}')\psi_n' + \varepsilon_n'\psi_n^0. \tag{31}$$

Substitution of Eqs. (27) and (30) obtained above with the use of first-order perturbation theory and the expansion

$$\psi_n'' = \sum_i b_i \psi_i^0, \tag{32}$$

as before, will allow Eq. (31) to be written as

$$\sum_i b_i(\varepsilon_i^0 - \varepsilon_n^0)\psi_i^0 = -\sum_j{}' \frac{\left\langle n\left|\hat{H}'\right|j\right\rangle\left\langle j\left|\hat{H}'\right|n\right\rangle}{\varepsilon_j^0 - \varepsilon_n^0}\psi_j^0 + \sum_j{}' \frac{\left\langle j\left|\hat{H}'\right|n\right\rangle}{\varepsilon_j^0 - \varepsilon_n^0}\hat{H}'\psi_j^0$$
$$+ \varepsilon_n''\psi_n^0 \ . \tag{33}$$

If this expression is multiplied by $\psi_n^{0\star}$ and integrated (with the condition that $n \neq j$), the result is

$$0 = 0 - \sum_j{}' \frac{\left\langle n\left|\hat{H}'\right|j\right\rangle\left\langle j\left|\hat{H}'\right|n\right\rangle}{\varepsilon_j^0 - \varepsilon_n^0} + \varepsilon_n'', \tag{34}$$

or

$$\varepsilon_n'' = \sum_j{}' \frac{\left\langle n\left|\hat{H}'\right|j\right\rangle\left\langle j\left|\hat{H}'\right|n\right\rangle}{\varepsilon_j^0 - \varepsilon_n^0}. \tag{35}$$

12.2.4 The anharmonic oscillator

The interatomic potential function for the diatomic molecule was described in Section 6.5. In the Taylor-series development of this function [Eq. (6-72)] cubic and higher terms were neglected in the harmonic approximation. It is now of interest to evaluate the importance of these so-called anharmonic terms with the aid of the perturbation theory outlined above. If cubic and quartic terms are added to Eq. (6-74), the potential function becomes

$$V(x) = \tfrac{1}{2}\kappa x^2 + ax^3 + bx^4 \tag{36}$$

where $x = r - r_e$. Clearly, the constants a and b are proportional to the equilibrium values of the third and fourth derivatives of the potential function, respectively.

With the Hamiltonian for the harmonic oscillator,

$$\hat{H}^0 = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \tfrac{1}{2}kx^2, \tag{37}$$

the Schrödinger equation leads to the solutions in Hermite polynomials (see Section 5.5.1). The zero-order wavefunctions are then

$$\psi_n^0 = \mathcal{N}_n \mathcal{H}_n(\xi) e^{-\frac{1}{2}\xi^2}, \tag{38}$$

where $\xi = 2\pi x\sqrt{\nu^0 m/\hbar} = \sqrt{\alpha}x$, $\varepsilon^0 = h\nu^0\left(n + \frac{1}{2}\right)$ and $\mathcal{N}_n$ is given by Eq. (5-111). The anharmonicity will now be considered as a perturbation, with $\hat{H}' = ax^3 + bx^4$. With the use of Eq. (27) the first-order correction to the energy is equal to

$$\left\langle n|\hat{H}'|n\right\rangle = a\left\langle n|x^3|n\right\rangle + b\left\langle n|x^4|n\right\rangle. \tag{39}$$

The evaluation of the matrix elements in Eq. (39) proceeds as follows.

First of all, consider the parity of the integrands. In the first term on the right-hand side of Eq. (39) both wavefunctions are either odd or even, thus their product is always even, while x^3 is of course odd. The integral between symmetric limits of the resulting odd function of x vanishes and this term makes no contribution to the first-order perturbation. On the other hand the second term is different from zero, as x^4 is an even function.

The matrix elements of x^4 can be evaluated with the use of the recursion relation developed in Section 5.5.1 for the Hermite polynomials (See Appendix IX). In the notation employed here Eq. (5-99) becomes

$$\xi\mathcal{H}_n(\xi) = \tfrac{1}{2}\mathcal{H}_{n+1}(\xi) + n\mathcal{H}_{n-1}(\xi). \tag{40}$$

Multiplication of Eq. (40) by ξ yields

$$\xi^2\mathcal{H}_n(\xi) = \tfrac{1}{2}\xi\mathcal{H}_{n+1}(\xi) + n\xi\mathcal{H}_{n-1}(\xi). \tag{41}$$

Equation (40) is then applied twice, *e.g.* with $n \to n + 1$ to give

$$\xi\mathcal{H}_{n+1}(\xi) = \tfrac{1}{2}\mathcal{H}_{n+2}(\xi) + (n+1)\mathcal{H}_n(\xi), \tag{42}$$

and similarly with $n \to n - 1$, to yield

$$\xi^2\mathcal{H}_n(\xi) = \tfrac{1}{4}\mathcal{H}_{n+2}(\xi) + \left(n + \tfrac{1}{2}\right)\mathcal{H}_n(\xi) + n(n-1)\mathcal{H}_{n-2}(\xi). \tag{43}$$

The square of Eq. (43) is given by

$$\xi^4 \mathcal{H}_n^2(\xi) = \tfrac{1}{16}\mathcal{H}_{n+2}^2(\xi) + \left(n + \tfrac{1}{2}\right)^2 \mathcal{H}_n^2(\xi) + n^2(n-1)^2 \mathcal{H}_{n-2}^2(\xi) + \mathcal{O}, \tag{44}$$

where $\mathcal{O}$ represents the cross terms which will vanish on integration due to the orthogonality of the wavefunctions. The second term on the right-hand side of Eq. (39) can be written in the form

$$b\langle n|x^4|n\rangle = \frac{b}{\alpha^2}\mathcal{N}_n^2 \int_{-\infty}^{\infty} e^{-\xi^2}\xi^4 \mathcal{H}_n^2(\xi)\,\mathrm{d}\xi. \tag{45}$$

With the substitution of Eq. (44) in the integrand, Eq. (45) becomes

$$b\langle n\left|x^4\right|n\rangle = \frac{b}{\alpha^2}\mathcal{N}_n^2\left[\frac{1}{16\mathcal{N}_{n+2}^2} + \frac{\left(n+\frac{1}{2}\right)^2}{\mathcal{N}_n^2} + \frac{n^2(n-1)^2}{\mathcal{N}_{n-2}^2}\right] \tag{46}$$

$$= \frac{3b}{2\alpha^2}\left[\left(n+\tfrac{1}{2}\right)^2 + \tfrac{1}{4}\right] \tag{47}$$

(problem 7). Thus, the first-order correction to the energy as a result of anharmonicity is proportional to the coefficient b and is quadratic in the vibrational quantum number n.

It was shown above that the cubic term in the potential function for the anharmonic oscillator cannot, for reasons of symmetry, contribute to a first-order perturbation. However, if the matrix elements of $\mathcal{H}' = ax^3$ are evaluated, it is found that this term results in a second-order correction to the energy. The appropriate matrix elements in Eq. (35) must then be evaluated.

The only nonvanishing matrix elements of x^3 are those with $j = n \pm 1$ and $j = n \pm 3$. This result is obtained by repeated application of Eq. (40), as before. Thus, there are four terms that the cubic potential constant contributes to the second-order energy correction, Eq. (35). The final result can be written as

$$\varepsilon_n'' = \frac{15a^2}{4\alpha^3 h\nu^0}\left(n^2 + n + \frac{11}{30}\right) = \frac{15a^2}{4\alpha^3 h\nu^0}\left[\left(n+\tfrac{1}{2}\right)^2 + \frac{17}{60}\right] \tag{48}$$

(problem 8). This expression, plus the first-order correction given by Eq. (47), indicates that the anharmonicity of the oscillator can be represented in this approximation by a quadratic term in the vibrational quantum number. The vibrational spectra of diatomic molecules are usually interpreted with the addition of an anharmonic term that is proportional to $\left(n+\frac{1}{2}\right)^2$. It should be noted that the application of perturbation theory yields, in addition, a small correction to the zero-point energy.

12.2.5 Degenerate systems

The first-order perturbation theory developed in Section 12.2.2 cannot be employed if the energy level of the unperturbed system is degenerate. In Eq. (12) it was assumed that the perturbed wavefunction ψ_n differed but slightly from a particular zero-order wavefunction, ψ_n^0. However, if the energy level ε_n^0 is α-fold degenerate, there are α linearly independent wavefunctions that satisfy the wave equation for the unperturbed system. Each of these functions is orthogonal to all wavefunctions corresponding to other energy levels. However, they are not necessarily orthogonal to each other.

The simplest example is that of a doubly degenerate level, for which $\hat{H}^0\psi_{k,1}^0 = \varepsilon_k^0\psi_{k,1}^0$ and $\hat{H}^0\psi_{k,2}^0 = \varepsilon_k^0\psi_{k,2}^0$. Clearly, any linear combination of the two wavefunctions is also a solution, as

$$\hat{H}^0(c_1\psi_{k,1}^0 + c_2\psi_{k,2}^0) = \varepsilon_k^0(c_1\psi_{k,1}^0 + c_2\psi_{k,2}^0). \tag{49}$$

In general, if α_k is the degree of degeneracy, α_k linear combinations of the zero-order wavefunctions can be constructed,

$$\chi_{k,i}^0 = \sum_{j=1}^{\alpha_k} c_{i,j}\psi_{k,j}^0 \qquad i = 1, 2, 3, \ldots, \alpha_k \ . \tag{50}$$

They are also correct wavefunctions for the zero-order problem. The coefficients can of course be chosen to normalize each result.

Consider a first-order perturbation. The Hamiltonian for the perturbed system is again as given by $\hat{H} = \hat{H}^0 + \lambda\hat{H}'$ [Eq. (12)], but the Schrödinger equation is of the form

$$\hat{H}\psi_{k,j} = \varepsilon_{k,j}\psi_{k,j} \ . \tag{51}$$

Thus, the effect of the perturbation may be to remove (all or partially) the degeneracies of the unperturbed energy levels. As the perturbation diminishes ($\lambda \to 0$), $\psi_{k,j} \to \chi_{k,i}^0$ and $\varepsilon_{k,j} \to \varepsilon_k^0$. Thus, for the perturbed system,

$$\psi_{k,i} = \chi_{k,i}^0 + \lambda\psi_{k,i}' \tag{52}$$

and

$$\varepsilon_{k,i} = \varepsilon_k^0 + \lambda\varepsilon_{k,i}' \ . \tag{53}$$

The first-order approximation (equating coefficients of λ) yields the relation

$$\hat{H}^0\psi_{k,i}' + \hat{H}'\chi_{k,i}^0 = \varepsilon_k^0\psi_{k,i}' + \varepsilon_{k,i}'\chi_{k,i}^0 \ . \tag{54}$$

Following the procedure of Section 12.2.2, the expansion

$$\psi'_{k,i} = \sum_{k,j} a_{k,i,k',j} \psi^0_{k',j} \tag{55}$$

is made in terms of the zero-order wavefunctions. Note that the double summation has been employed in Eq. (55) to take into account all wavefunctions associated with a given unperturbed energy level, as well as all energy levels.

Substitution of Eqs. (52) and (53) in Eq. (51) leads to

$$\sum_{k',j} a_{k,i,k',j}(\varepsilon^0_{k'} - \varepsilon^0_k)\psi'_{k',j} = \sum_j c_{i,j}(\varepsilon'_{i,j} - \hat{H}')\psi^0_{k,j} \ , \tag{56}$$

which, when multiplied by $\psi^{0^*}_{k,n}$ and integrated over all space becomes

$$0 = \sum_{j=1}^{\alpha_k} c_{i,j} \left[\varepsilon'_{k,i} \int \psi^{0\bigstar}_{k,n} \psi^0_{k,j} \, \mathrm{d}\tau - \int \psi^{0\bigstar}_{k,n} \hat{H}' \psi^0_{k,j} \, \mathrm{d}\tau \right] \tag{57}$$

or

$$\sum_{j=1}^{\alpha_k} c_{i,j} \left[\left\langle n \left| \hat{H}' \right| j \right\rangle - \varepsilon'_{k,i} \langle n | j \rangle \right] = 0, \qquad i = 1, 2, 3, \ldots, \alpha_k \ . \tag{58}$$

This result is a system of simultaneous linear, homogeneous equations for the coefficients, $c_{i,j}$. Cramer's rule states that a nontrivial solution exists only if the determinant of the coefficients vanishes (see Section 7.8). Thus,

$$\left| \left\langle n \left| \hat{H}' \right| j \right\rangle - \varepsilon'_{k,i} \langle n | j \rangle \right| = 0, \qquad i = 1, 2, 3, \ldots, \alpha_k \ . \tag{59}$$

The determinant in Eq. (59) is of course a secular determinant, a description that refers to its application to the temporal evolution of a mechanical system, historically in astronomy. It will re-appear later in this chapter in the development of the variation method.

The secular determinant as presented above involves the first-order perturbations of the Hamiltonian and the energy. More generally, it is formulated in terms of the Hamiltonian and the total energies of the perturbed system. From Eqs. (12) and (16),

$$\left\langle n | \hat{H}' | j \right\rangle = \left\langle n | \hat{H}^0 | j \right\rangle + \left\langle n | \hat{H}' | j \right\rangle \tag{60}$$

and

$$\varepsilon_{k,i} \langle n | j \rangle = \varepsilon^0_k \langle n | j \rangle + \varepsilon'_{k,i} \langle n | j \rangle \ . \tag{61}$$

Their difference, when substituted in Eq. (58) leads to the secular determinant in its more usual form, *viz.*

$$\left|\left\langle n\left|\hat{H}\right|j\right\rangle-\varepsilon_{k,i}\left\langle n|j\right\rangle\right|=0,\qquad i=1,2,3,\ldots,\alpha_k. \tag{62}$$

12.2.6 The Stark effect of the hydrogen atom

Consider a hydrogen atom under the influence of an electric field, $\mathcal{E}$. The perturbation energy is given by $\hat{H}'=-\boldsymbol{\mu}\cdot\mathcal{E}$, where $\boldsymbol{\mu}$ is the instantaneous dipole moment of the atom. For simplicity, assume that the electric field is directed along the z axis. The perturbation in this case then given by

$$\hat{H}'=ez\mathcal{E}_z=e\mathcal{E}_z r\cos\theta, \tag{63}$$

where e is the electronic charge.

The ground state of the hydrogen atom is nondegenerate and the wavefunction is

$$\psi_{1,0,0}=\frac{1}{\sqrt{4\pi}}R_{1,0}, \tag{64}$$

with the radial part given by

$$R_{1,0}=2a_0^{-3/2}e^{-r/a_0}. \tag{65}$$

The result of first-order perturbation theory, as given by Eq. (27), is applicable. Clearly, the matrix element $\left\langle 1,0,0\left|\hat{H}'\right|1,0,0\right\rangle$ vanishes, as $\int_0^{\pi}\cos\theta\sin\theta\, d\theta=0$. Thus, there is no first-order Stark effect for the hydrogen atom in the ground state.

However, in the first excited state the degree of degeneracy is equal to four. Hence, the first-order perturbation calculation requires the application of Eq. (62). The wavefunctions for the first excited state can be written in the form

$$\begin{aligned}
\psi_{2,0,0}&=\frac{1}{\sqrt{4\pi}}R_{2,0}\\
\psi_{2,1,0}&=\sqrt{\frac{3}{4\pi}}R_{2,1}\cos\theta\\
\psi_{1,1,1}&=\sqrt{\frac{3}{4\pi}}R_{2,1}\sin\theta\frac{1}{\sqrt{2}}e^{i\varphi}\\
\psi_{1,1,-1}&=-\sqrt{\frac{3}{4\pi}}R_{2,1}\sin\theta\frac{1}{\sqrt{2}}e^{-i\varphi},
\end{aligned} \tag{66}$$

where the radial parts are given by

$$R_{2,0} = \frac{a_0^{-3/2}}{2\sqrt{2}}\left(2 - \frac{r}{a_0}\right)e^{-r/2_0^a}$$

$$R_{2,1} = \frac{a_0^{-3/2}}{2\sqrt{6}}\frac{r}{a_0}e^{-r/2_0^a} \tag{67}$$

In Eqs. (67) the quantity a_0 is the radius of the first Bohr orbit (see Section 6.6).

The matrix elements in Eq. (62) are of two types. The diagonal elements involve integrals over θ, all of which vanish. Furthermore, of the off-diagonal elements, only $\left\langle 2, 0, 0 \middle| \hat{H}' \middle| 2, 1, 0 \right\rangle = \left\langle 2, 1, 0 \middle| \hat{H}' \middle| 2, 0, 0 \right\rangle = -3e\mathcal{E}_z a_0$ are nonzero, as all integrals involving φ vanish. The secular determinant then takes the form

$$\begin{vmatrix} -\varepsilon_2' & -3e\mathcal{E}_z a_0 & 0 & 0 & 0 \\ -3e\mathcal{E}_z a_0 & -\varepsilon_2' & 0 & 0 & 0 \\ 0 & 0 & 0 & -\varepsilon_2' & 0 \\ 0 & 0 & 0 & 0 & -\varepsilon_2' \end{vmatrix} = 0 \tag{68}$$

and the energies of the perturbed levels are obtained as the roots of the expression

$$\varepsilon_2'^2(\varepsilon_2'^2 - 9e^2\mathcal{E}_z^2 a_0^2) = 0 \tag{69}$$

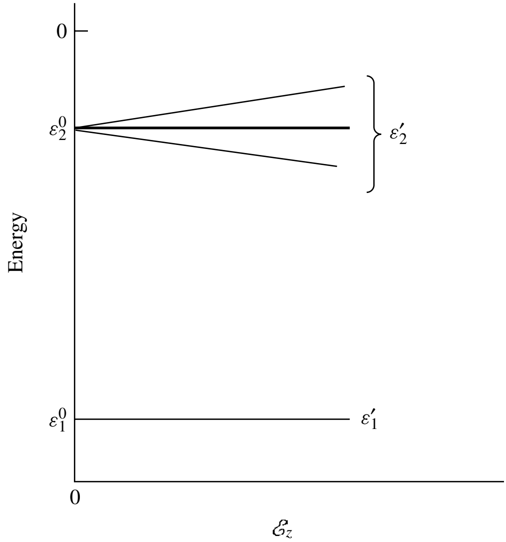


Fig. 1 The Stark effect of the hydrogen atom.

(problem 11). These results are represent in Fig. 1. Note that two of the levels are not perturbed by the applied electric field and their degeneracy is not removed. They correspond to atomic orbitals p_x and p_y, whose maxima lie in the horizontal plane, that is, perpendicular to the direction of the applied field (see Appendix III).

12.3 TIME-DEPENDENT PERTURBATIONS

12.3.1 The Schrödinger equation

To discuss the problem of the interaction of light with an atomic or molecular system, it is essential to consider those perturbations which are functions of time. It is such perturbations that provoke transitions from one stationary state to another in a given system. Thus, specifically, it is the time-dependent Schrödinger equation that must be considered. It can be written in the form

$$\hat{H}\Psi(q_1, q_2, \cdots; t) = i\hbar \frac{\partial \Psi(q_1, q_2, \cdots; t)}{\partial t}, \tag{70}$$

where the Hamiltonian is now a function of time, as well as all of the coordinates of the system.

The separation of the time from the spatial variables can be carried out by the method introduced in Section 6.1.2. Thus, it is sufficient to write

$$\Psi(q_1, q_2, \cdots; t) = \psi(q_1, q_2, \cdots)\vartheta(t) \tag{71}$$

to separate the time from the space variables. The result is given by

$$\Psi(q_1, q_2, \cdots; t) = \psi(q_1, q_2, \cdots)e^{-i\varepsilon_k t/\hbar}. \tag{72}$$

In the problem of interest here, the Hamiltonian in Eq. (62) can be decomposed into a time-independent, unperturbed part $\hat{H}^0$ and a much smaller, time-dependent operator $\hat{H}'(t)$. Then, the Hamiltonian becomes to first order

$$\hat{H}^0 + \hat{H}'(t). \tag{73}$$

The stationary states of the system are described by the eigenfunctions ψ_k and the eigenvalues ε_k of the unperturbed Hamiltonian.

The eigenfunctions ψ_k form a complete set; thus, the wavefunction $\Psi(q_1, q_2, \cdots; t)$ can be expanded in terms of the wavefunctions ψ_k, with the use of the time-dependent coefficients $b_k(t)$. The resulting expression is then

$$\Psi(q_1, q_2, \cdots; t) = \sum_k b_k(t)\psi_k e^{-i\varepsilon_k t/\hbar} \tag{74}$$

which can be substituted into the time-dependent Schrödinger equation to yield

$$\sum_k b_k(t)\hat{H}'(t)\psi_k e^{-i\varepsilon_k t/\hbar} = i\hbar \sum_k \frac{\mathrm{d}b_k(t)}{\mathrm{d}t}\psi_k e^{-i\varepsilon_k t/\hbar}. \tag{75}$$

Multiplication by $\psi_m^\star$ and integration over the spatial coordinates results in the set of differential equations for the coefficients,

$$\frac{\mathrm{d}b_m(t)}{\mathrm{d}t} = -\frac{i}{\hbar}\sum_k b_k(t)\left\langle m \left| \hat{H}'(t) \right| k \right\rangle e^{-(\varepsilon_k - \varepsilon_m)t/\hbar}. \tag{76}$$

It should be noted that in the general case each term in the summation contains three time-dependent factors. Thus, to reduce the complexity of the presentation, only a very simple example will be considered here.

Suppose that at a time $t = 0$, the stationary state of the system is known. That is to say that both its energy ε_n and the corresponding (nondegenerate) wavefunction ψ_n have been determined. Therefore, for $t \leq 0$, all of the coefficients in Eq. (76) are equal to zero, except of course, $b_n = 1$, which identifies the initial state of the system. It is often the ground state, although not necessarily. A perturbation is then applied at $t = 0$. It is then assumed that it is sufficiently weak so that the coefficient b_n does not vary significantly from its initial value of unity over the relatively short duration of the perturbation. With these conditions Eq. (76) can be reduced to

$$\frac{\mathrm{d}b_m(t)}{\mathrm{d}t} = -\frac{i}{\hbar}\left\langle m \left| \hat{H}'(t) \right| n \right\rangle e^{-(\varepsilon_n - \varepsilon_m)t/\hbar}. \tag{77}$$

Normally, the energy difference between the initial and final states, as given in Eq. (77), is equated to $\omega_{nm}\hbar$, where $\nu_{nm} = \omega_{nm}/2\pi$ is the frequency of the transition from the initial state $k = n$ to the final state m.

12.3.2 Interaction of light and matter

The problem of particular interest in physics and chemistry is concerned with the interaction of electromagnetic radiation, and light in particular, with matter. The electric field of the radiation can directly perturb an atomic or molecular system. Then, as in the Stark effect, the energy of interaction – the perturbation – is given by

$$\hat{H}' = -\boldsymbol{\mu}\cdot\mathcal{E}, \tag{78}$$

where $\boldsymbol{\mu}$ is the dipole moment of the system. For simplicity it will be assumed that the light incident on the system is polarized in a particular direction, say x. Then, Eq. (78) becomes simply

$$\hat{H}' = -\mu_x\mathcal{E}_x\ . \tag{79}$$

If the incident radiation is monochromatic with a frequency $\nu = \omega/2\pi$, as in a laser beam, its electric field can be represented by

$$\mathcal{E}_x = \mathcal{E}_x^0 \, cos \, \omega t, \tag{80}$$

where $\mathcal{E}_x^0$ is the amplitude of the incident radiation. The corresponding component (x) of the dipole moment of the system can be written in the form

$$\mu_x = \sum_i e_i x_i \ , \tag{81}$$

where e_i is the charge on each particle, and x_i is its position in the x direction. Then the perturbation is given by

$$\hat{H}' = -\mu_x \mathcal{E}_x = -\mathcal{E}_x \sum_i e_i x_i \tag{82}$$

and the matrix elements in Eq. (77) are

$$\left\langle m \left| \hat{H}'(t) \right| n \right\rangle = -\mathcal{E}_x^0 \, cos \, \omega t \sum_i e_i \left\langle m \left| \mu_i \right| n \right\rangle = -\mathcal{E}_x^0 \, cos \, \omega t \left\langle m \left| \mu_x \right| n \right\rangle . \tag{83}$$

The time dependence is in this case due to the oscillation of the imposed electric field as given by Eq. (80), as well as the displacement of the charged particles, electrons and nuclei within the atomic or molecular system.

The expression for the matrix elements given by Eq. (83) is substituted in Eq. (77). The result is

$$\frac{\mathrm{d}b_m(t)}{\mathrm{d}t} = \frac{i}{\hbar} \mathcal{E}_x^0 \, cos \, \omega t \left\langle m \left| \mu_x \right| n \right\rangle e^{-i\omega_{nm} t}. \tag{84}$$

The exponential form of $cos \, \omega t$, as given in Eq. (1-36), is then substituted to obtain

$$\frac{\mathrm{d}b_m(t)}{\mathrm{d}t} = \frac{i}{2\hbar} \mathcal{E}_x^0 \left\langle m \left| \mu_x \right| n \right\rangle [e^{-i(\omega_{nm} - \omega)t} + e^{-i(\omega_{nm} + \omega)t}] \tag{85}$$

which can be easily integrated if the matrix element $\langle m | \mu_x | n \rangle$ does not change significantly during the short time of the perturbation. Then, with the initial condition that $b_m = 0$ at $t = 0$, integration yields the time-dependent coefficient

$$b_m(t) = \frac{-\mathcal{E}_x^0 \left\langle m \left| \mu_x \right| n \right\rangle}{2\hbar} \left[\frac{e^{-i(\omega_{nm} - \omega)t} - 1}{\omega_{nm} - \omega} + \frac{e^{-i(\omega_{nm} - \omega)t} - 1}{\omega_{nm} + \omega} \right]. \tag{86}$$

As the frequency of the incident electromagnetic radiation approaches that of the transition $n \rightarrow m$, the first term in the brackets of Eq. (86) dominates; it becomes very large at resonance.* Thus, the second term in Eq. (79) can be neglected in the spectral region of interest, and the resulting expression for the probability of a transition to the state m can be written as (see problems 14 and 15)

$$b_m b_n^* = \frac{\left|\mathscr{E}_x^0\right|^2 |\langle m\,|\mu_x|\,n\rangle|^2\, t^2}{4\hbar^2}\left\{\frac{sin^2[(\omega_{nm}-\omega)t/2]}{[(\omega_{nm}-\omega)t/2]^2}\right\}. \tag{87}$$

The expression in brackets in Eq. (87) is of the form $(sin\,x/x)^2$, where $x = (\omega_{nm} - \omega)t/2$. Thus, for a given time t for the duration of the perturbation, the spectrum, *e.g.* the transition probability as a function of the angular frequency ω is as shown in Fig. 2. The width at half-maximum of this spectral feature is represented by Δ for a given value of the time, t. If, for example, the perturbation time is increased by a factor of four, the width of the spectral distribution is reduced by the same factor, as shown by the solid line in Fig. 2. Equation (87) expresses the probability that the system, initially in the state $k = n$, will be in the state m after a sinusoidal perturbation over a relatively short period of time t.

Now to calculate the transition probability in the case in which all frequencies $\nu = \omega/2\pi$ are incident on the molecule, Eq. (87) must be integrated over the frequency range. As the significant frequency variation is due to the factor

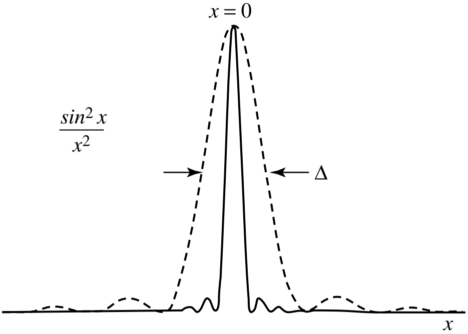


Fig. 2 The function $(sin\,x/x)^2 = sinc^2 x$, where $x = (\omega_{nm} - \omega)t/2$ and t is the duration of the perturbation. The width of the frequency distribution is equal to Δ, which is proportional to $1/t$.

---

*Note, however, that it does not become infinite, as $lim_{a\to 0}[(e^{iat} - 1)/a] = it$.

in brackets, its integration leads to the approximate expression

$$|b_m|^2 = \frac{\pi \left|\mathscr{E}_x^0\right|^2}{2\hbar^2} \left|\langle m\,|\mu_x|\,n\rangle\right|^2 t. \tag{88}$$

The quantity $|b_m|^2$ represents the probability of the transition $m \leftarrow n$. Clearly, the number of transitions per unit time depends on the intensity of the incident radiation, which is proportional to $|\mathscr{E}_x^0|^2$, and the square of the matrix element $\langle m|\mu_x|n\rangle$. The latter determines the selection rules for spectroscopic transitions (see the following section).

The result obtained as Eq. (88) can be generalized for the case of isotropic radiation. The light intensity is then proportional to

$$|\mathscr{E}^0|^2 = |\mathscr{E}_x^0|^2 + |\mathscr{E}_y^0|^2 + |\mathscr{E}_z^0|^2, \tag{89}$$

which is in turn proportional to the radiation density, ρ. The square of the matrix elements of the dipole-moment vector are given by

$$|\,\langle m|\boldsymbol{\mu}|n\rangle\,|^2 = |\,\langle m|\mu_x|n\rangle\,|^2 + |\left\langle m|\mu_y|n\right\rangle|^2 + |\,\langle m|\mu_z|n\rangle\,|^2, \tag{90}$$

which depends on the molecular orientation. It determines the value of the Einstein coefficient for absorption, which is given by

$$B_{m\leftarrow n} = \frac{2\pi}{(4\pi\varepsilon_0)3c\hbar^2}|\,\langle m|\boldsymbol{\mu}|n\rangle\,|^2. \tag{91}$$

The quantity $\langle m|\boldsymbol{\mu}|n\rangle$ is known as the transition (dipole) moment.

In the above rather simplified analysis of the interaction of light and matter, it was assumed that the process involved was the absorption of light due to a transition $m \leftarrow n$. However, the same result is obtained for the case of light emission stimulated by the electromagnetic radiation, which is the result of a transition $m \rightarrow n$. Then the Einstein coefficients for absorption and stimulated emission are identical, *viz.* $B_{m\leftarrow n} = B_{m\rightarrow n}$.

An important process has not been included in the analysis. It is the possibility of spontaneous emission. Were it not for such a process, in the absence of electromagnetic radiation a molecule in the excited state m would be forced to remain there forever. Thus, in Einstein's analysis of this problem three competing processes were considered to be in equilibrium, leading to the expression

$$B_{m\leftarrow n}N_n\rho = B_{m\rightarrow n}N_m\rho + A_{m\rightarrow n}N_m\ . \tag{92}$$

The left-hand side of Eq. (92) represents the number of transitions per unit time due to absorption of light. The first term on the right-hand side of Eq. (92)

is due to stimulated emission, while the second term, which remains in the absence of electromagnetic radiation, takes into account the possibility of spontaneous emission. As an equation, this relation expresses the balance between these processes. Furthermore, at equilibrium the relative populations of the upper and lower states is determined by the Boltzmann distribution law,

$$\frac{N_m}{N_n} = e^{-\hbar\omega_{nm}/kT} \tag{93}$$

(see Chapter 10). Substitution of Eq. (93) into Eq. (92) yields the relations between the Einstein coefficients

$$A_{m\to n} = \frac{2\hbar\omega^3}{\pi c^2} B_{m\to n} = \frac{2\hbar\omega^3}{\pi c^2} B_{m\leftarrow n}\ . \tag{94}$$

It is important to note that all three coefficients depend on the matrix elements of the dipole moment, as expressed by Eq. (90).

12.3.3 Spectroscopic selection rules

General selection rules that govern spectroscopic transitions are derived from the symmetry properties of the dipole moment and the wavefunctions involved. The transition moments can be expressed in general by

$$\langle m|\mu_j|n\rangle = \int \psi_m^\star \mu_j \psi_n \,\mathrm{d}\tau, \tag{95}$$

where $j = x, y, z$ identifies space-fixed Cartesian coordinates and ψ_m and ψ_n are the wavefunctions for the states involved in the transition. If any such integral is nonzero, from Eq. (90) it is evident that these transitions are possible. The so-called selection rules are just the answer to the question: Is the transition moment $\langle m|\boldsymbol{\mu}|n\rangle$ equal to zero? If so, transitions between the states m and n are forbidden.

The determination of general selection rules can be made by consideration of the symmetry of the integrand in Eq. (95). For example, it was shown in Section 3.4.5 that the integral over an odd function vanishes. The corresponding group-theoretical expression of this principle was outlined in Section 8.10. The integrand in Eq. (95) is the product of three functions. Thus, if each is characterized by a representation, the direct product provides the needed information concerning the symmetry of the resulting function. In the simplest case each factor may belong to a particular irreducible representation. Then, the resulting direct product can be expressed as

$$\Gamma_m^{(\gamma')} \otimes \Gamma_j^{(\gamma)} \otimes \Gamma_n^{(\gamma'')} = \sum_\ell n^{(\ell)}\Gamma^{(\ell)}. \tag{96}$$

The right-hand side of Eq. (96) is of course the weighted direct sum of the irreducible representations. By convention the totally symmetric irreducible representation corresponds to $\ell = 1$. Thus, if $n^{(1)} = 0$, the integral in Eq. (95) vanishes. The transitions $m \to n$ and $m \leftarrow n$ are then forbidden by the symmetry selection rules. This principle can be illustrated by the following example.

Consider a molecular system of symmetry $\mathcal{C}_{2\mathrm{v}}$, whose character table is given in Table 8-11. The irreducible representations for the components of the dipole moment can be easily established, or even read directly from the table. Thus $\Gamma_j^{(\gamma)}$ can be identified as $\mathrm{A}_1, \mathrm{B}_1$ or B_2 for $j = z, x$ or y, respectively. Then, if the direct product $\Gamma_m^{(\gamma')} \otimes \Gamma_n^{(\gamma'')}$ contains any of these three irreducible representations, the transitions $m \to n$ and $m \leftarrow n$ are allowed. Furthermore, the polarization of optical transitions can be specified, as each $\Gamma_j^{(\gamma)}$ corresponds to a specific polarization direction.

The selection rules illustrated above are general, as they depend only on the symmetry properties of the functions involved. However, more limiting, selection rules depend on the form of the wavefunctions involved. A relatively simple example of the development of specific selection rules is provided by the harmonic oscillator. The solution of this problem in quantum mechanics, as treated in Section 5.4.4, leads to the wavefunctions given by Eq. (5-103) and the energy levels defined by Eq. (5-92). These results were employed in Section 5.4.4 to describe in a first approximation the vibration of a diatomic molecule. As before, the possibility of transitions between the various energy levels of the system is determined by the matrix elements of the dipole moment.

If r is the internuclear distance in a diatomic molecule and $x = r - r_e$, the dipole moment can be developed in a series in the form

$$\mu = \mu_0 + \left(\frac{\mathrm{d}\mu}{\mathrm{d}x}\right)_0 x + \dots \; . \tag{97}$$

The first term on the right-hand side of Eq. (97) is the permanent dipole moment. The second term expresses the change in dipole moment with internuclear distance. Often, higher terms are neglected. The derivative $(\mathrm{d}\mu/\mathrm{d}x)_0$ can then be interpreted as an effective charge carried by the vibrating nuclei. The change in internuclear distance is related to the independent variable employed in Section 5.5.1 by $\xi = 2\pi x\sqrt{\nu^0 m/h}$, where m is now the reduced mass of the diatomic molecule and $\nu^0 = \sqrt{k/m}/2\pi$ is the classical frequency of vibration.

The transition moment of interest for process $v' \to v$, is then

$$\langle v'|\mu|v\rangle = \left\langle v' \middle| \mu_0 + \left(\frac{\mathrm{d}\mu}{\mathrm{d}x}\right)_0 x \middle| v \right\rangle = \mu_0 \langle v'|v\rangle + \left(\frac{\mathrm{d}\mu}{\mathrm{d}x}\right)_0 \langle v'|x|v\rangle . \tag{98}$$

For a polar molecule ($\mu_0 \neq 0$) the first term on the far right is nonzero only if the initial and final vibrational states are the same, *viz.* $v = v'$. This case applies to the pure rotational spectra of gaseous molecules, as observed in the microwave region. The second term in Eq. (98) applies to vibrational transitions. The matrix elements of interest are $\langle v'|x|v\rangle$, which are given by

$$\begin{aligned}\langle v'|x|v\rangle &= \alpha^{-1/2}\langle v'|\xi|v\rangle \\ &= \alpha^{-1/2}\mathcal{N}_{v'}\mathcal{N}_v \int_{-\infty}^{+\infty} (e^{-\xi^2/2}\mathcal{H}_{v'}(\xi))\xi(e^{-\xi^2/2}\mathcal{H}_v(\xi))\,\mathrm{d}\xi. \qquad (99)\end{aligned}$$

To determine the selection rules in this case it is sufficient to recall the relations developed in Section 5.5.1 between the Hermite polynomials. Specifically, Eq. (5-99) can be rewritten in the form

$$\xi\mathcal{H}_v(\xi) = \tfrac{1}{2}\mathcal{H}_{v+1}(\xi) + v\mathcal{H}_{v-1}(\xi). \qquad (100)$$

With its substitution in Eq. (99) it becomes evident from the orthogonality of the Hermite polynomials, that all matrix elements are equal to zero, with the exception of $v' = v - 1$ and $v' = v + 1$. Thus, the selection rule for vibrational transitions (in the harmonic approximation) is $\Delta v = \pm 1$. It is not necessary to evaluate the matrix elements unless there is an interest in calculating the intensities of spectral features resulting from vibrational transitions (see problem 18). It should be evident that transitions such as $\Delta v = \pm 3$ are forbidden under this more restrictive selection rule, although they are permitted under the symmetry selection rule developed in the previous paragraphs.

As a second example of the determination of selection rules from the properties of special functions, consider the hydrogen atom. At any given instant the dipole moment is $\boldsymbol{\mu} = e\boldsymbol{r}$, where $\boldsymbol{r}$ describes the position of the electron with respect to the proton and e is the electronic charge. The wavefunctions for the hydrogen atom are given by

$$\psi_{n,\ell,m} = R_{n,\ell}(r)\Theta_{\ell,m}(cos\,\theta)e^{im\varphi}, \qquad (101)$$

where the angular-dependent factor is given by Eq. (6-69). If the incident light is polarized in, say, the z direction, the matrix elements of interest are of the form

$$\begin{aligned}\langle n',\ell',m'|\mu_z|n,\ell,m\rangle &= e\langle n',\ell',m'|r\,cos\,\theta|n,\ell,m\rangle \\ &= e\langle n',\ell'|r|n,\ell\rangle\langle\ell',m'|cos\,\theta|\ell,m\rangle\langle m'|m\rangle. (102)\end{aligned}$$

The factor that depends on the radial wavefunctions is in general nonzero. The factor in φ contributes to the integral

$$\langle m'|m\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{-i(m'-m)}\,\mathrm{d}\varphi = \delta_{m',m}\ , \qquad (103)$$

which imposes the condition $\Delta m = 0$, or the integral vanishes. The θ-dependent part of Eq. (102) can be evaluated from the recursion relations for the associated Legendre polynomials. For this example Eq. (5-122) can be written as

$$cos\,\theta\, P_\ell^{|m|}(cos\,\theta) = \frac{(\ell + |m|)}{(2\ell + 1)} P_{\ell-1}^{|m|}(cos\,\theta) + \frac{(\ell - |m| + 1)}{(2\ell + 1)} P_{\ell+1}^{|m|}(cos\,\theta), \tag{104}$$

which leads directly to the selection rule $\Delta\ell = \pm\, 1$.

For light polarized in the x or y directions, the procedure followed in the preceding paragraph can be employed with $\mu_x = ex = er\,sin\,\theta\,cos\,\varphi$ and $\mu_y = ey = er\,sin\,\theta\,sin\,\varphi$. Then, it is apparent that $\Delta m = \pm 1$ and with the aid of Eqs. (5-123) and (5-124) the selection rule $\Delta\ell = \pm 1$ can be easily established (see problem 19). In conclusion the selection rules for the absorption of unpolarized light are $\Delta m = 0, \pm 1$ and $\Delta\ell = \pm 1$.

12.4 THE VARIATION METHOD

A different approach to obtaining approximation solutions to quantum mechanical problems is provided by the variation method. It is particularly useful when there is no closely related problem that yields exact solutions. The perturbation method is not applicable in such a case.

The variation method is usually employed to determine an approximate value of the lowest energy state (the ground state) of a given atomic or molecular system. It can, furthermore, be extended to the calculation of energy levels of excited states. It forms the basis of molecular orbital theory and that which is often referred to (incorrectly) as “theoretical chemistry”.

12.4.1 The variation theorem

Given an acceptable, normalized function ϕ, if the lowest eigenvalue of the Hamiltonian $\hat{H}$ is ε_0, then

$$W = \int \phi^\star \hat{H} \phi \, \mathrm{d}\tau \geq \varepsilon_0, \tag{105}$$

which is the variation theorem. This relation may be a bit surprising, as the function ϕ can be any normalized function of the coordinates of the system that satisfies the conditions for an acceptable wavefunction. Although the function ϕ is arbitrary, the more wisely it is chosen the more closely will W approach ε_0, the true energy of the ground state of the system. Thus, if ϕ were chosen to be the correct ground-state wavefunction, ψ_0, Eq. (105) would yield the energy of the ground state.

As a simple proof of the variational theorem, consider the case in which $\phi \neq \psi_0$. The variational function can be expanded in terms of the complete set of normalized, orthogonal functions ψ_n. Thus,

$$\phi = \sum_n a_n \psi_n \ , \tag{106}$$

with

$$\sum_n a_n^{\star} a_n = 1. \tag{107}$$

Substitution of this expansion in the integral for W [Eq. (103)] leads to the equation

$$W = \sum_n \sum_{n'} a_n^{\star} a_{n'} \int \psi_n^{\star} \hat{H} \psi_{n'} \, \mathrm{d}\tau = \sum_n a_n^{\star} a_n \varepsilon_n \ , \tag{108}$$

as the functions ψ_n satisfy the equation

$$\hat{H} \psi_n = \varepsilon_n \psi_n \ . \tag{109}$$

The energy of the ground state, ε_0 is then subtracted from each side of Eq. (108) to yield

$$W - \varepsilon_0 = \sum_n a_n^{\star} a_n (\varepsilon_n - \varepsilon_0), \tag{110}$$

where Eq. (107) has been employed.

As ε_n is greater than ε_0 for all values of n, and the coefficients $a_n^{\star} a_n$ are of course positive or zero, the right-hand side of Eq. (108) is positive or zero. It has thus been shown that W is always an upper limit to ε_0, the true energy of the ground state of the system; thus,

$$W \geq \varepsilon_0. \tag{111}$$

Equation (111) is a statement of the variational theorem.

If several variation functions, $\phi_1, \phi_2, \phi_3, \ldots$ are chosen and the corresponding values of the variational energy, W_1, W_2, $W_3, \ldots$ are calculated from Eq. (103), each of these values of W will be greater than the true energy of the ground state, ε_0. Thus, the lowest one is nearest ε_0. In many cases it is convenient to employ a variational function that contains one or more parameters. Then the resulting expression for W can be minimized with respect to the parameters.

12.4.2 An example: The particle in a box

The one-dimensional problem of the particle in a box was treated in Section 5.4.1. Exact solutions were obtained, which were then restricted by the boundary conditions $\psi(0) = \psi(\ell) = 0$. If the exact solutions were not known, the problem might be attacked with the use of a simple variation function which satisfies the boundary conditions. As an example, take

$$\phi = \mathcal{N} x(\ell - x), \tag{112}$$

which vanishes at each side of the box. The normalization constant can be evaluated from the relation

$$\int_0^\ell \phi^2 \, d\phi = \mathcal{N}^2 \int_0^\ell x^2(\ell - x)^2 \, dx = \mathcal{N}^2 \ell^5/30, \tag{113}$$

which leads to $\mathcal{N} = \sqrt{30/\ell^5}$ (problem 21).

In the interior of the box $V(x) = 0$ and the Hamiltonian is simply

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}; \tag{114}$$

then, Eq. (105) becomes

$$W = -\frac{h^2}{8\pi^2 m}\frac{30}{\ell^5}\int_0^\ell x(\ell - x)\frac{d^2}{dx^2}[x(\ell - x)] \, dx = \frac{5h^2}{4\pi^2 m\ell^2}. \tag{115}$$

The true energy of the ground state was found in Section 5.4.1 to be equal to $\varepsilon_1 = h/8m\ell^2$. Its comparison with Eq. (108) is

$$W = \frac{10}{\pi^2}\left(\frac{h^2}{8m\ell^2}\right) > \frac{h^2}{8m\ell^2} = \varepsilon_1 \tag{116}$$

and the error is in this case approximately 2%.

Although the variational theorem was expressed in Eq. (111) with respect to the ground state of the system, it is possible to apply it to higher, so-called excited, states. As an example, consider again the particle in a box. In Section 5.4.2 a change in coordinate was made in order to apply symmetry considerations. Thus, the potential function was written as

$$V(x) = \begin{cases} 0, & -\frac{1}{2}\ell < x < \frac{1}{2}\ell \\ \infty, & x = -\frac{1}{2}\ell, \frac{1}{2}\ell \end{cases}. \tag{117}$$

It is an even function of x and the solutions can be classified as even or odd (Gerade or Ungerade), as given in Eqs. (5-72) and (5-73). It should be noted that the ground state in this case is symmetric. It is then of interest to choose a variation function that is antisymmetric (u) to determine the energy of the first excited state $n = 2$, which is antisymmetric. As an example, consider the variational function

$$\phi = \mathcal{N} x(\ell/2 + x)(\ell/2 - x) = \mathcal{N} x\left[(\ell^2/4) - x^2\right], \tag{118}$$

which is Ungerade. The normalization leads to the factor $\mathcal{N} = \sqrt{840/\ell^7}$ (see problem 21). The variational energy is then given by

$$W = \frac{42}{\pi^2}\left(\frac{h^2}{8m\ell^2}\right) > 4\left(\frac{h^2}{8m\ell^2}\right), \tag{119}$$

where on the right-hand side the factor $4 = n^2$ specifies the first excited state, $n = 2$.

12.4.3 Linear variation functions

It is often convenient to employ a variation function that it is a linear combination of suitably chosen functions χ_n; thus,

$$\phi = \sum_n^m c_n \chi_n. \tag{120}$$

It should be noted that the functions χ_n need not necessarily form an orthonormal set. The linearly independent coefficients c_n can be considered to be variable parameters that are determined by minimization of the variational energy, W. If the functions χ_n are not orthonormal, Eq. (105) can be rewritten in the form

$$W = \frac{\int \phi^\star \hat{H} \phi \, \mathrm{d}\tau}{\int \phi^\star \phi \, \mathrm{d}\tau} \geq \varepsilon_0. \tag{121}$$

The variational energy is then given by

$$W = \frac{\sum_n^m \sum_n^m c_{n'}^\star c_n \left\langle n' \middle| \hat{H} \middle| n \right\rangle}{\sum_n^m \sum_{n'}^m c_{n'}^\star c_n \left\langle n' | n \right\rangle}, \tag{122}$$

where $\langle n' | \hat{H} | n \rangle \equiv \int \chi_{n'}^\star \hat{H} \chi_n \, \mathrm{d}\tau$ and $\langle n' | n \rangle \equiv \int \chi_{n'}^\star \chi_n \, \mathrm{d}\tau$. The partial derivative of Eq. (122) with respect to a particular coefficient $c_k^\star$ leads to the relation

$$\begin{aligned} &\frac{\partial W}{\partial c_k^\star} \sum_{n'}^m \sum_n^m c_{n'}^\star c_n \langle n' | n \rangle + W \frac{\partial}{\partial c_k^\star}\left(\sum_{n'}^m \sum_n^m c_{n'}^\star c_n \langle n' | n \rangle\right) \\ &\quad = \frac{\partial}{\partial c_k^\star}\left(\sum_{n'}^m \sum_n^m c_{n'}^\star c_n \left\langle n' \middle| \hat{H} \middle| n \right\rangle\right). \end{aligned} \tag{123}$$

If the coefficients are independent, the condition $\partial W/\partial c_k^\star = 0$ can be imposed for each value of the index k from 1 to m.* Then,

$$W\sum_{n}^{m} c_n \langle k|n\rangle = \sum_{n}^{m} c_n \left\langle k|\hat{H}|n\right\rangle, \tag{124}$$

or,

$$\sum_{n}^{m} c_n \left(\left\langle k\left|\hat{H}\right|n\right\rangle - W\langle k|n\rangle\right) = 0. \tag{125}$$

Equation (125) applies for all values of the index $k = 1, 2, \ldots, m$. It is a set of m simultaneous, homogeneous, linear equations for the unknown values of the coefficients c_n. Following Cramer's rule (Section 7.8), a nontrivial solution exists only if the determinant of the coefficients vanishes. Thus, the secular determinant takes the form

$$|\left\langle k|\hat{H}|n\right\rangle - W\langle k|n\rangle| = 0. \tag{126}$$

In the case in which the functions χ_n are orthonormal, $\langle k|n\rangle = \delta_{k,n}$ and the variational energies W are just the eigenvalues of the matrix $\langle k|\hat{H}|n\rangle$. According to the variational theorem, the lowest root of Eq. (124) is the upper limit to the true energy of the ground state of the system, ε_0.

12.4.4 Linear combinations of atomic orbitals (LCAO)

It is often convenient to use atomic orbitals as the basis for molecular-orbital calculations. Thus, in Eq. (120) the atomic orbitals χ_n can serve as the basis, and a given molecular system can be described as a linear combination of such functions. Clearly, the simplest molecule is diatomic and the appropriate molecular orbitals can be formed as linear combinations, *viz.*

$$\phi = c_a\chi_a + c_b\chi_b, \tag{127}$$

where the functions χ_a and χ_b are the atomic orbitals associated with the atoms a and b, respectively. The coefficients in this linear combination can be determined by application of the variational principle, as illustrated in the following derivation. This method, which is of general application to polyatomic molecules, is referred to in the scientific literature as the method LCAO.

---

*Note that the equivalent condition $\partial W/\partial c_k = 0$ yields a set of equations which is simply the complex conjugate of Eq. (122).

A given application of the LCAO method is characterized by a set of integrals. For a diatomic molecule they are

$$H_{aa} = \left\langle \chi_a \left| \hat{H} \right| \chi_a \right\rangle, \tag{128}$$

$$H_{bb} = \left\langle \chi_b \left| \hat{H} \right| \chi_b \right\rangle, \tag{129}$$

$$H_{ab} = \left\langle \chi_a \left| \hat{H} \right| \chi_b \right\rangle = \left\langle \chi_b \left| \hat{H} \right| \chi_a \right\rangle \tag{130}$$

and

$$S = \langle \chi_a | \chi_b \rangle\,. \tag{131}$$

The last is known as the overlap integral, as it is determined by the volume common to the atomic orbitals a and b at a given internuclear distance. In general, $S < 1$, an integral that is often set equal to zero in approximate calculations.

The use of a linear variation function was summarized in the previous section. For the example of a diatomic molecule the set of simultaneous equations [Eq. (125)] becomes

$$c_a(H_{aa} - W) + c_b(H_{ab} - SW) = 0 \tag{132}$$

and

$$c_a(H_{ab} - SW) + c_b(H_{bb} - W) = 0. \tag{133}$$

The expansion of the corresponding secular determinant leads to the relation

$$(H_{aa} - W)(H_{bb} - W) - (H_{ab} - SW)^2 = 0, \tag{134}$$

which is quadratic in W. To evaluate the coefficients c_a and c_b the two values of W are substituted successively in either Eq. (132) or Eq. (133), as described in the classical example of Section 7.11. However, as these equations are homogeneous, only the ratio of the coefficients can be determined. The supplementary condition necessary to resolve this ratio is provided by normalization of the functions ϕ. The atomic orbitals are assumed to be normalized, or can be made so. Thus, normalization of the molecular orbital in this case can be expressed by

$$\langle \phi | \phi \rangle = c_a^2 + c_b^2 + 2c_a c_b S = 1. \tag{135}$$

This example of the LCAO method, as applied to diatomic molecules, is perfectly general. However, it is simpler for homonuclear diatomics, for which $H_{aa} = H_{bb}$. Then, Eq. (134) becomes

$$(H_{aa} - W)^2 - (H_{ab} - SW)^2 = 0, \tag{136}$$

whose roots are

$$W_{\pm} = \frac{H_{aa} \pm H_{ab}}{1 \pm S}. \tag{137}$$

Thus, two values can be evaluated, W_+ and W_-, according to the signs in Eq. (137). If the appropriate integrals are known, these quantities can be calculated for a given interatomic distance in the diatomic molecule. As indicated above the successive substitution of the two values W_+ and W_- yields

$$c_a \mp c_b = 0. \tag{138}$$

With the application of the normalization condition given by Eq. (135), the coefficients in Eq. (138) are found as

$$c_a = c_b = \frac{1}{\sqrt{2(1+S)}} \tag{139}$$

and

$$c_a = -c_b = \frac{1}{\sqrt{2(1-S)}}, \tag{140}$$

respectively, depending on whether the upper or lower sign is employed in the preceding equations. The variational wavefunctions are then of the form

$$\phi_+ = \frac{1}{\sqrt{2(1+S)}}(\chi_a + \chi_b) \tag{141}$$

and

$$\phi_- = \frac{1}{\sqrt{2(1-S)}}(\chi_a - \chi_b), \tag{142}$$

with the corresponding energies given by Eq. (137).

The simplest diatomic species is the molecular ion H_2^+. Its electronic kinetic energy is given by a single term, as that of the protons can be neglected (see Section 12.1). The interaction of the electron with each proton is expressed by Coulomb's law, as is the proton–proton repulsion. With the use of the resulting Hamiltonian the integrals defined by Eqs. (128)–(130) can be evaluated exactly, with the functions χ_a and χ_b the 1s orbitals of atomic hydrogen (see Section 6.6). The resulting energies calculated with the use of Eq. (137) are represented in Fig. 3, as it is the lowest energy level of H_2^+ that is of interest. Clearly, the curve of W_+ exhibits a minimum and, if the corresponding state ϕ_+ is "occupied" by the electron, a stable species can exist. This orbital is then referred to as a bonding orbital. On the other hand the curve of W_- *vs.* internuclear distance has no minimum, so its occupation by the electron cannot

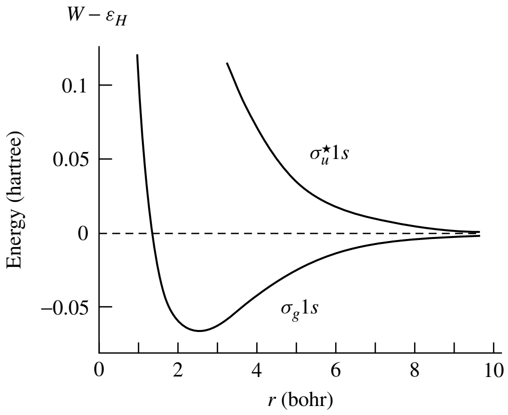


Fig. 3 Variational energy of H_2^+ as a function of internuclear distance.

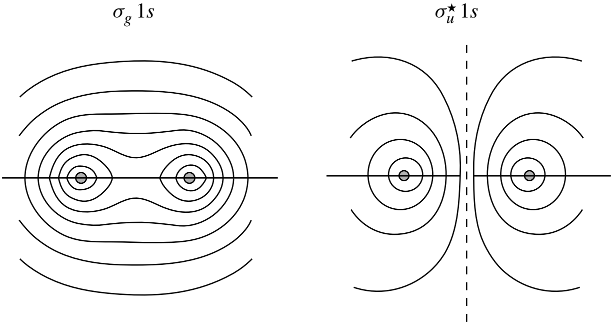


Fig. 4 The wavefunctions $\phi_+(\sigma_g 1s)$ and $\phi_-(\sigma_u^\bigstar 1s)$ for H_2^+. Note that the wavefunction ϕ_- has opposite signs on either side of the dotted line.

result in bonding of the two protons. Thus, the orbital ϕ_- is an antibonding orbital and, if occupied by the electron, the ion is in a so-called excited state.

It should be noted (Fig. 4) that if the bonding orbital, designated σ_g 1s, is occupied, the probability of finding the electron in the region between the nuclei is relatively important. On the other hand, for the antibonding orbital ($\sigma_u^\bigstar 1s$) there is a nodal plane that passes through the center of symmetry of the ion. It is perpendicular to the internuclear axis. Thus, it can be concluded that it is the electronic probability density that is responsible for bonding. Although this conclusion is correct and can be generalized, the wavefunctions obtained by the LCAO method are usually far from the true functions. Furthermore, although the lower energy level as calculated by the method is an upper limit to the true energy of the ground state, the variation theorem in its simplest form says nothing about the value of the upper energy level.

The limitation of the above analysis to the case of homonuclear diatomic molecules was made by imposing the relation $H_{aa} = H_{bb}$, as in this case the two nuclei are identical. More generally, $H_{aa} \neq H_{bb}$, and for heteronuclear diatomic molecules Eq. (134) cannot be simplified (see problem 25). However, the polarity of the bond can be estimated in this case. The reader is referred to specialized texts on molecular orbital theory for a development of this application.

12.4.5 The Hückel approximation*

One of the most popular of the semi-empirical LCAO methods is that of Hückel. It is applicable to planar molecules which have π-electron systems. The "delocalization" of these systems, as treated by this method, has particular chemical significance. The traditional application is to the benzene molecule. Historically, different "structures" of this molecule were suggested by Kékulé† and by Dewar‡ that are described in virtually all textbooks of organic chemistry. These structures represented the first efforts to represent the delocalization of the π orbitals in such systems. In the present context the delocalization can be better specified with the use of the method of Hückel.

Consider first the ethylene molecule. Its geometrical structure is shown in Fig. 5. The s, p_y and p_z atomic orbitals of the carbon atoms are assumed to be hybridized. This sp^2 hybridization implies H–C–H bond angles of 120°, approximately in agreement with experimental results. The remaining two p_x orbitals are thus available to contribute to a π-electron system in the molecule. Here again, the two linear combinations of atomic orbitals yield bonding and

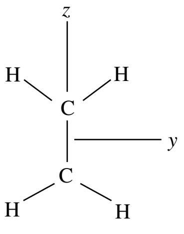


Fig. 5 The ethylene molecule showing only the single (σ) bonds.

---

*Erich Hückel, German chemist (1896–1980).

†August Kékulé von Stradonitz, German chemist (1829–1896).

‡Sir James Dewar, British chemist and physicist (1842–1923).

antibonding possibilities. The LCAO method presented in the previous section can be employed to obtain a semiquantitative description of the electronic structure.

For a homonuclear diatomic system in the Hückel approximation the integrals given by Eqs. (128)–(131) take the simple forms $H_{aa} = H_{bb} = \alpha$, $H_{ab} = H_{ba} = \beta$ and $S = 0$. The atomic orbitals involved, χ_a and χ_b, are of course the p_x orbitals of carbon atoms a and b, respectively. The resulting secular determinant is then simply

$$\begin{vmatrix} \alpha - W & \beta \\ \beta & \alpha - W \end{vmatrix} = 0, \tag{143}$$

which can be written as

$$\begin{vmatrix} x & 1 \\ 1 & x \end{vmatrix} = 0, \tag{144}$$

with $x = (\alpha - W)/\beta$. As the roots of Eq. (144) are $x = \pm 1$, the energy levels are determined by $W = \pm(\alpha - \beta)$, as shown in Fig. 6. The corresponding approximate molecular orbitals are also indicated, with the normalizing factor $1/\sqrt{2}$. The coefficients were evaluated as illustrated in the previous section [Eqs. (139) and (140)]. Here again, the occupied lower level will result in an increased electronic density in the region between the nuclei, although the upper level, if occupied, will not.

As the ethylene molecule contains a total of 16 electrons, there are but two that are available to occupy the π system. Two pairs of electrons are assumed to fill the two 1s atomic orbitals of the carbon atoms. Five pairs of electrons contribute to the σ orbitals that represent single bonds in Fig. 5. Thus, the two

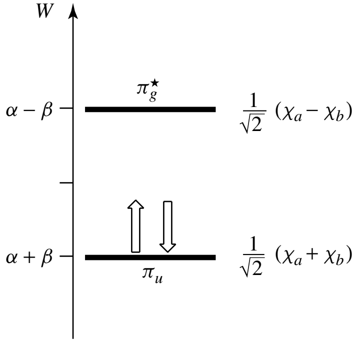


Fig. 6 The π energy levels of ethylene. The asterisk identifies the antibonding orbital, while the two arrows represent the two electrons with antiparallel spins corresponding to the configuration of the ground state (see text).

remaining can occupy the molecular orbitals established by the LCAO method, as outlined above. For the ground state of the molecule these two electrons, with antiparallel spins following the Pauli principle, occupy the lower energy level, as shown in Fig. 6. Excited states can be described by promoting one or both electrons to the higher energy level.

The ethylene molecule in its equilibrium configuration is of symmetry $\mathscr{D}_{2\mathrm{h}}$. Its symmetry can then be used to simplify the development of the appropriate molecular orbitals. Thus, the characters of the reducible representation of the π orbitals can be determined, as given in Table 1. With the application of the magic formula (see Section 8.9) it is easy to establish the reduction of the representation for the two π orbitals, namely, $\mathbf{\Gamma}_\pi = \mathrm{B}_{2\mathrm{g}} \oplus \mathrm{B}_{3\mathrm{u}}$. The appropriate linear combination of atomic orbitals can then be found with the use of the projection operator technique. However, as only two π orbitals are involved, it is sufficient to consider a subgroup of $\mathscr{D}_{2\mathrm{h}}$ that is of order two and that includes an operation that exchanges the labels (a,b) on the atomic orbitals. The group $\mathscr{S}_2 \equiv \mathscr{C}_i$ is appropriate, as it preserves the g–u property of the molecular orbitals. The characters for this group, as given in Appendix VIII, are in this case just the coefficients of the atomic wavefunctions, χ_a and χ_b. The notation for the irreducible representations shown in Fig. 6 is identified in Table 2.

In conjugated systems the π orbitals become delocalized. The classical example is the butadiene molecule, that is usually described by the formula $CH_2{=}CH{-}CH{=}CH_2$. This representation of the molecule does not take into consideration the delocalization of the π-electron system formed by the four

Table 1 Symmetry operations of the group $\mathscr{D}_{2\mathrm{h}}$ and the characters of $\mathbf{\Gamma}_\pi$.

$\mathscr{D}_{2\mathrm{h}}$	E	$C_2(z)$	$C_2(y)$	$C_2(x)$	i	$\sigma(xy)$	$\sigma(xz)$	$\sigma(yz)$
$\boldsymbol{\chi}_\pi$	2	−2	0	0	0	0	2	−2

Table 2 The character table for the group $\mathscr{S}_2 \equiv \mathscr{C}_i$ and the characters of $\mathbf{\Gamma}_\pi$.

$\mathscr{S}_2 \equiv \mathscr{C}_i$	E	i
$\mathrm{A_g} \equiv \pi_\mathrm{g}^*$	1	1
$\mathrm{A_u} \equiv \pi_\mathrm{u}$	1	−1
$\boldsymbol{\chi}_\pi$	2	0

available p orbitals of the carbon atoms. To treat this problem with the use of the method of Hückel, it is sufficient to write the secular determinant as a function of the integrals α and β and to introduce the approximations indicated above. In addition to the often poor approximation $S = 0$, it is customary to assume that there is no interaction between nonadjacent atomic orbitals. Thus, the integrals of the type H_{ab} are set equal to zero if the atoms a and b are not adjacent. The butadiene molecule provides a good example of this method.

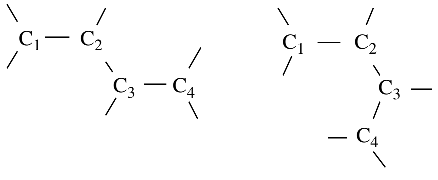
Butadiene exists in two equilibrium structural isomers. They are represented in Fig. 7. However, with the usual Hückel approximation these two structures cannot be distinguished, as interactions between nonadjacent atoms have been neglected. Thus for either isomer, or even a hypothetical structure in which the carbon skeleton is linear, the secular determinant is the same, namely,

$$\begin{vmatrix} x & 1 & 0 & 0 \\ 1 & x & 1 & 0 \\ 0 & 1 & x & 1 \\ 0 & 0 & 1 & x \end{vmatrix} = 0, \tag{145}$$

where $x = (\alpha - W)/\beta$, as before. The roots of this equation can be found by direct expansion of the determinant (see problem 26). However, the application of group theory allows the determinant to be factored, a technique that is often useful in more complicated problems.

Consider the *trans* isomer of butadiene. Both the symmetry operations that define the group $\mathscr{C}_{2\mathrm{h}}$ and the characters of the representation $\mathbf{\Gamma}_\pi$ are given in Table 3. The reduction of this representation leads to $\mathbf{\Gamma}_\pi = 2\mathrm{B_g} \oplus 2\mathrm{A_u}$. Thus, two linear combinations of the atomic orbitals can be constructed of symmetry $\mathrm{B_g}$ and two others of symmetry $\mathrm{A_u}$. Their use will factor the secular determinant into two 2×2 blocks, as described in the following paragraph.

As indicated above, the conventional structure of *trans* butadiene does not include the delocalization of the π-electron system. This effect can be analyzed, at least approximately, by application of the Hückel method. As in the example of ethylene, each carbon atom has an available p orbital–the



(**a**) (**b**)

Fig. 7 The equilibrium structures of butadiene: *trans* (**a**) and *cis* (**b**). Only the σ bonds are shown.

Table 3 The symmetry operations of the group $\mathcal{C}_{2h}$ and the characters of $\boldsymbol{\Gamma}_\pi$.

$\mathcal{C}_{2h}$	E	C_2	i	σ_h
$\boldsymbol{\chi}_\pi$	4	0	0	−4

p_z orbitals in this case. With the use of the projection operator, the linear combinations of the four atomic orbitals can be constructed, *viz.*

$$B_g \begin{cases} \frac{1}{\sqrt{2}}(\chi_1 - \chi_4) \\ \frac{1}{\sqrt{2}}(\chi_2 - \chi_3) \end{cases} \tag{146}$$

and

$$A_u \begin{cases} \frac{1}{\sqrt{2}}(\chi_1 + \chi_4) \\ \frac{1}{\sqrt{2}}(\chi_2 + \chi_3) \end{cases} \tag{147}$$

where the atomic numbering is as shown in Fig. 7. The two diagonal blocks of the factored secular determinant are then of the form

$$B_g : \begin{vmatrix} x & 1 \\ 1 & x-1 \end{vmatrix} = 0 \tag{148}$$

and

$$A_u : \begin{vmatrix} x & 1 \\ 1 & x+1 \end{vmatrix} = 0. \tag{149}$$

The roots of Eqs. (148) and (149) are given by $x = (1 \pm \sqrt{5})/2$ and $x = (-1 \pm \sqrt{5})/2$, respectively.* The results of the above analysis are summarized in Fig. 8.

The benzene molecule in its equilibrium configuration is planar. Its symmetry is described by the point group $\mathcal{D}_{6h}$ as shown in Fig. 8-1(c). The delocalized π system is represented there by dotted lines. The six p_z orbitals contribute to the π system, as simply described by the Hückel approximation. The reduction $\boldsymbol{\Gamma}_\pi = B_{2g} \oplus E_{1g} \oplus A_{2u} \oplus E_{2u}$ can be found as in the previous examples. However, to construct the appropriate linear combinations of the π orbitals, it is sufficient to choose a subgroup of $\mathcal{D}_{6h}$ whose symmetry operations permute all

---

*The quantity $(1+\sqrt{5})/2 = 1.62$ is known as the golden ratio. It appears often in works of art, as for example to determine the approximate ratio of the height to width of a classic painting – and this page.

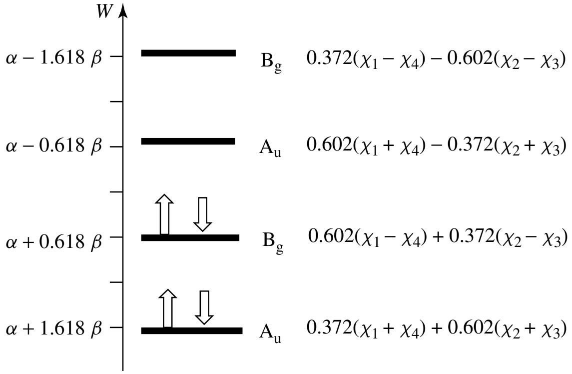


Fig. 8 The energy levels and wavefunctions for the π-orbitals of *trans* butadiene. The arrows define the electron configuration of the ground state.

six of the p_z atomic orbitals. For example, the choice of the group $\mathscr{C}_6$ leads to the equivalent reduction $\mathbf{\Gamma}_\pi = \mathrm{A} \oplus \mathrm{B} \oplus \mathrm{E}_1 \oplus \mathrm{E}_2$. The character table for this group (see Appendix VII) contains complex elements of the type $\varepsilon = exp(2\pi i/6)$. The application of the projection operator will then yield linear combinations such as

$$\mathrm{E}_1 : \begin{cases} \chi_1 + \varepsilon\chi_2 - \varepsilon^\star\chi_3 - \chi_4 - \varepsilon\chi_5 + \varepsilon^\star\chi_6 \\ \chi_1 + \varepsilon^\star\chi_2 - \varepsilon\chi_3 - \chi_4 - \varepsilon^\star\chi_5 + \varepsilon\chi_6 \end{cases}. \tag{150}$$

As $\varepsilon + \varepsilon^\star = 1$ in this case, the equivalent linear combinations that involve only real coefficients can be found by adding and subtracting the two functions given in Eq. (150). For the π system of benzene the normalized linear combinations of p_z can then be determined, as given in Fig. 9. It is easily verified that these linear combinations are orthonormal. Furthermore, they result in the desired factoring of the secular determinant. The construction of the energy level diagram is then relatively straightforward (problem 29).

The use of group theory to factor the secular determinant is of increasing importance as the molecule becomes larger, providing of course that it maintains a relatively high symmetry. With the use of available computer programs the advantages of this approach may seem to be of less interest. However, it should be understood that symmetry arguments lead to visualization of the molecular orbitals, as now represented by a number of programs. Their comprehension requires the basic understanding of the elements of group theory, as outlined above.

As a final exercise for the reader, consider the naphthalene molecule (symmetry $\mathscr{D}_{2h}$), as shown in Fig. 10. Application of the Hückel method leads to a 10×10 secular determinant (see problem 30). However, with the application

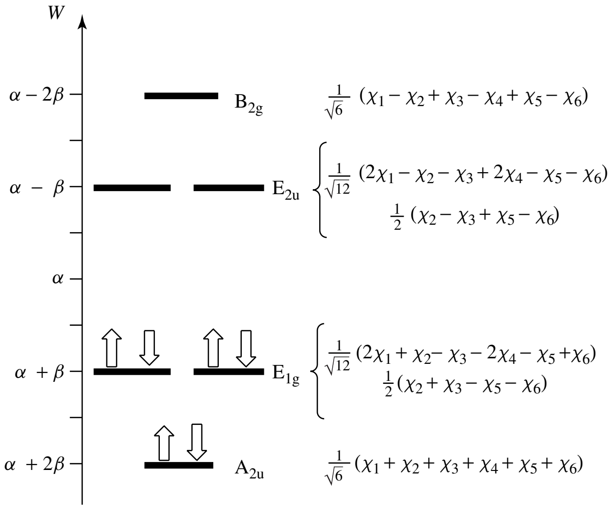


Fig. 9 The energy levels and the LCAO orbitals for the π-electron system of benzene. The electron configuration as represented by the arrows is that of the ground state of the molecule.

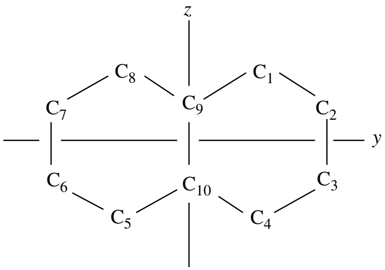


Fig. 10 The naphthalene molecule. Only the σ bonds are represented.

of the method outlined above the reduced representation takes the form $\mathbf{\Gamma}_\pi = 2B_{1g} \oplus 3B_{2g} \oplus 2A_u \oplus 3B_{3u}$. The appropriate linear combination of p_x orbitals can be constructed directly with the use of the projection operator.

PROBLEMS

1. Verify Eq. (10).
2. Develop the series of relations given by Eqs. (18)–(...).

3. Derive Eq. (26).

4. Derive Eq. (30).

5. Derive Eq. (35).

6. Make the indicated substitutions to obtain Eq. (44).

7. Verify Eq. (47).

8. Derive Eq. (48).

9. Derive the secular equations, Eq. (58).

10. Verify Eq. (62).

11. Verify Eq. (69).

12. Carry out the separation of variables to obtain Eq. (72).

13. Derive Eq. (76).

14. Make the indicated substitution to obtain Eq. (85).

15. Derive Eqs. (86) and (87).

16. Integrate Eq. (87) as indicated.

17. Derive the relation between the Einstein coefficients [Eq. (94)].

18. Evaluate the transition moments in Eq. (99) for $v' = v \pm 1$.

Ans. See Appendix IX.

19. Verify the selection rules for the hydrogen atom as given in the last paragraph of Section 12.3.3.

20. Prove the variation theorem [Eq. (111)].

21. Calculate the normalization factor in Eqs. (113) and (118).

22. Verify Eq. (115).

23. Verify Eq. (119).

24. Derive the secular determinant [Eq. (126)].

25. Show that the variational energies of a homonuclear diatomic molecule are given in the LCAO approximation by Eq. (137) and that the corresponding wavefunctions are as indicated in Eqs. (141) and (142).

26. Find the roots of Eq. (145) by direct expansion of the determinant and compare them with the results given below, Eq. (149).

27. Show that $\mathbf{\Gamma}_\pi = 2\mathrm{B}_\mathrm{g} \oplus 2\mathrm{A}_\mathrm{u}$ for the π-electron system in *trans* butadiene.

28. Verify Eqs. (146) and (147).

29. Apply the projection-operator method to obtain the molecular orbital expression shown in Fig. 9 and verify the energies.

30. Set up the secular determinant for the π-system of naphthalene and factor it as explained in the last paragraph of this chapter.

13 Numerical Analysis

13.1 ERRORS

All measurements are subject to errors. It is therefore essential in any scientific endeavor to analyze the results of experiments and to estimate the reliability of the data obtained. In general, experimental errors can be classified as

(i) systematic,
(ii) personal, and
(iii) random.

Systematic errors can arise in measurements made with a given scientific instrument. Often the acquired data do not represent directly the quantity of interest. Furthermore, the instrument used may not be correctly calibrated. These, and other sources of error that are inherent in an experiment, give rise to systematic errors. With care on the part of the experimentalist they can often be detected and, one hopes, corrected.

Personal errors sometimes occur due to inattention, or even prejudice, on the part of an observer. Sometimes an experimentalist would very much like to obtain data that agree with his hypothesis. His reaction – that is, to cheat slightly – may be entirely subjective. Furthermore, he may quite unconsciously make mistakes, either in his observation or in the subsequent presentation of his results. A well known example of the latter type of error was in an early report of the concentration of iron in spinach. According to that communication, spinach was found to be an incredibly rich source of iron. This result was propagated in the literature – including the image created by Popeye to encourage the young to eat more spinach. It was not a bad idea, of course, but a decimal-point error in the early experimental results was responsible for the exaggeration of the iron content of spinach.

If systematic errors can be traced, and perhaps eliminated, and personal errors can be minimized, the remaining random errors can be analyzed by statistical methods. This procedure will be summarized in the following sections.

13.1.1 The Gaussian distribution

Consider the probability

$$\mathscr{W}(x) = p^x q^{n-x} C(n, x) = \frac{p^x q^{n-x} n!}{(n-x)!x!}, \tag{1}$$

as given by Eq. (10-14) for the Bernoulli trials. Out of n trials, p is the number of successes and q the number of failures. When n is large, Eq. (1) can be approximated by a Gaussian distribution. This result is obtained by taking the logarithm and substituting Stirling's approximate expression [Eq. (10-21)] for each factorial. Then,

$$ln\, \mathscr{W}(x) = x\, ln\, \frac{np}{x} + (n-x)\, ln\, \frac{nq}{n-x} + \tfrac{1}{2}\, ln\, \frac{n}{2\pi x(n-x)}. \tag{2}$$

The logarithms in the first two terms in Eq. (2) can be developed in a power series, as shown in Section 2.9, *e.g.*

$$ln\, \frac{np}{x} = -\, ln\left(1 + \frac{x-np}{np}\right) = -\left[\frac{x-np}{np} - \tfrac{1}{2}\left(\frac{x-np}{np}\right)^2 + \cdots\right], \tag{3}$$

and similarly for the logarithm of $nq/(n-x)$. Then, as $p + q = 1$,

$$\mathscr{W}(x) \approx \frac{1}{\sqrt{2\pi npq}} e^{-(x-np)^2/2npq}, \tag{4}$$

if only the first term in each series is retained (problem 2).

The function $\mathscr{W}(x)$ is that of Gauss, which was discussed in Section 3.4.5. It is presented in Fig. 3-4, although the normalization condition is in this case somewhat different. As $\mathscr{W}(x)\,\mathrm{d}x$ represents a probability, its integration over all of the sample space must yield the certainty. The function is thus normalized in the sense that

$$\int_{-\infty}^{\infty} \mathscr{W}(x)\, \mathrm{d}x = 1. \tag{5}$$

The approximations introduced above are quite satisfactory close to the origin, although they become questionable further away. However, it is just in the latter regions that the exponential becomes weak. Thus, for most practical purposes Eq. (4) is a good approximation to the probability distribution, as the number of samples becomes large.

It is customary to define the dispersion of the distribution by

$$\sigma^2 = npq. \tag{6}$$

The quantity σ is known as the standard deviation. Furthermore, the mean value of the random variable x is given by $np = \bar{x}$. Then, Eq. (4) becomes

$$\mathscr{W}(x) \approx \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\bar{x})^2/2\sigma^2}. \tag{7}$$

The result obtained here has particular significance in the analysis of random errors of measurement. The substitution $t = (x - \bar{x})/\sigma$ in Eq. (7) leads to the expression

$$\mathscr{W}(t) \approx \frac{1}{\sqrt{2\pi}\sigma} e^{-t^2/2}, \tag{8}$$

which is Gauss's error function. From Eq. (7) it can be concluded that the probability that a given measurement yields a value of x in the interval $\pm x$ is given by

$$\mathscr{W}(x) \approx \frac{1}{\sqrt{2\pi}} \int_{-x}^{+x} e^{-t^2/2}\,\mathrm{d}t = \frac{1}{\sqrt{\pi}} \int_{-x}^{+x} e^{-y^2}\,\mathrm{d}y, \tag{9}$$

which can be written as

$$\mathit{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{+x} e^{-y^2}\,\mathrm{d}y. \tag{10}$$

The integral in Eq. (10) is the usual definition of the error function. A closely related function is the complementary error function

$$\mathit{erfc}(x) = 1 - \mathit{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-y^2}\,\mathrm{d}y. \tag{11}$$

The error function cannot be evaluated analytically, although it is readily available in the form of tables and evaluated in many computer programs.

13.1.2 The Poisson distribution*

In the previous section it was assumed that quantities of the order of $1/np$ and $1/npq$ were negligible. In that case the mean value of np is a large number. However, in many applications the quantity p is small and the product np remains finite. In this case the distribution is spread out, although the mean value remains small. The resulting distribution is no longer symmetrical. This behavior is illustrated in Fig. 1.

---

*Siméon Denis Poisson, French mathematician (1781–1840).

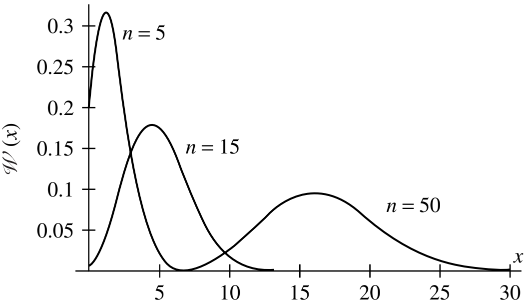


Fig. 1 The Poisson distribution [Eq. (13)] with $p = \frac{1}{3}$.

Return now to the binomial distribution [Eq. (1)] and let n approach infinity. The result is then

$$\mathscr{W}_n(x) = \frac{[1 - (np/n)]^n (np)^x}{x!} \frac{1[1 - (1/n)] \cdots [1 - (x-1)/n]}{[1 - (np/n)]^x}; \tag{12}$$

thus,

$$\mathscr{W}(x) = \lim_{n \to \infty} \mathscr{W}_n(x) = \frac{(np)^x e^{-np}}{x!}. \tag{13}$$

Note that because the product np remains finite, the second factor in Eq. (12) approaches unity in the limit. Similarly, from the definition of the exponential (Section 1.4), $lim_{n\to\infty}[1 - (np/n)]^n = e^{-np}$. Equation (13) is an expression of the Poisson distribution.

The Poisson distribution is usually applied in the case of small values of np. For large values it is well approximated by the normal, or Gaussian, distribution. For a given value of p the distribution becomes more nearly symmetric with increasing values of n. It becomes wider and approaches a Gaussian form, as shown in Fig. 1. This distribution, and others, are often approximated by the normal (Gaussian) distribution in the region near the maximum. Although there are many applications of the Poisson distribution, the best known is in the area of atomic physics. The result of counting particles emitted by a radioactive substance is usually described by the Poisson distribution.

13.2 THE METHOD OF LEAST SQUARES

The normal distribution, as expressed by Eq. (7), can be employed in the analysis of random errors. If the error in a given measurement i is represented by x_i, the probability that it lies between x_i and $x_i + \mathrm{d}x_i$ can be written as

$$P_i = \frac{1}{\sigma\sqrt{\pi}} e^{-x_i^2/\sigma^2}\, \mathrm{d}x_i \ . \tag{14}$$

Hence, for n independent measurements the combined probability is given by

$$P = \prod_{i=1}^{n} P_i = \left[\frac{1}{\sigma\sqrt{\pi}}\right]^n exp\left[-(1/\sigma^2)\sum_{i=1}^{n} x_i^2\right] \mathrm{d}x_1\, \mathrm{d}x_2 \ldots \mathrm{d}x_n \ . \tag{15}$$

For a given value of σ the probability is maximum when the sum in the exponent is minimum. Thus, the minimization of $\sum_{i=1}^{n} x_i^2$ becomes the criterion for the most probable value obtainable from n equally reliable measurements. This result is the basis for the various curve-fitting procedures that are commonly used in the analysis of experimental data.

It is very often of interest to fit a set of data points to a straight line. While it is possible to draw a line on a graph by eye, it is clearly preferable to have an objective method to establish the line with respect to the experimental points. Suppose that the straight line is specified by

$$Y_i = mx_i + b, \tag{16}$$

where m is its slope and b its intercept on the ordinate axis. The deviation of each point from the line is equal to $y_i - Y_i = \varepsilon_i$. The sum of squares, which is then given by

$$S = \sum_{i=1}^{n} (y_i - Y_i)^2 = \sum_{i=1}^{n} (y_i - mx_i - b)^2, \tag{17}$$

is the quantity to be minimized with respect to the two parameters m and b. Thus,

$$\frac{\partial S}{\partial b} = -2\sum_{i=1}^{n} (y_i - mx_i - b) = 0 \tag{18}$$

and

$$\frac{\partial S}{\partial m} = -2\sum_{i=1}^{n} (y_i - mx_i - b)x_i = 0. \tag{19}$$

The resulting expressions for the two parameters can be expressed as a function of the averages $\bar{x} = 1/n \sum_{i=1}^{n} x_i$, $\bar{y} = 1/n \sum_{i=1}^{n} y_i$, $\overline{x^2} = 1/n \sum_{i=1}^{n} x_i^2$ and $\overline{xy} = 1/n \sum_{i=1}^{n} x_i y_i$ in the form

$$m = \frac{\overline{x}\,\overline{y} - \overline{xy}}{\overline{x}^2 - \overline{x^2}} \tag{20}$$

and

$$b = \overline{y} - m\overline{x}. \tag{21}$$

Thus, the straight line corresponding to the best fit is established without ambiguity.

Many computer programs exist to achieve the linear, least-squares fitting (linear regression) to a given set of data. It is, however, worthwhile to apply the method to a simple problem in order to understand the basis. The data presented in Fig. 1-1 represent the weight of Miss X as a function of the date. While there is no reason to suppose that there is a linear relationship implied, the straight line provides her with an indication of her rate of weight loss, namely, the slope of the line. The least-squares fit to the data yields the relation $Y = -0.12x + 69.6$, as shown in the figure.

13.3 POLYNOMIAL INTERPOLATION AND SMOOTHING

Consider the simplest method of interpolating between two successive data points. It is linear, midpoint interpolation. This procedure is illustrated in Fig. 2. The ordinate value of the interpolated point is given by

$$Y = \tfrac{1}{2}y_{-1/2} + \tfrac{1}{2}y_{+1/2}, \tag{22}$$

the average of the values at the two points. The slope of the line segment connecting the two points is easily found as

$$\frac{\mathrm{d}Y}{\mathrm{d}x} = -y_{-1/2} + y_{+1/2}, \tag{23}$$

where the interval Δx has been taken equal to one. The coefficients appearing in Eqs. (22) and (23), when arranged in matrix form yield

$$\boldsymbol{A} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -1 & 1 \end{pmatrix}, \tag{24}$$

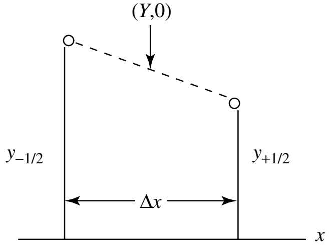


Fig. 2 Linear, midpoint interpolation.

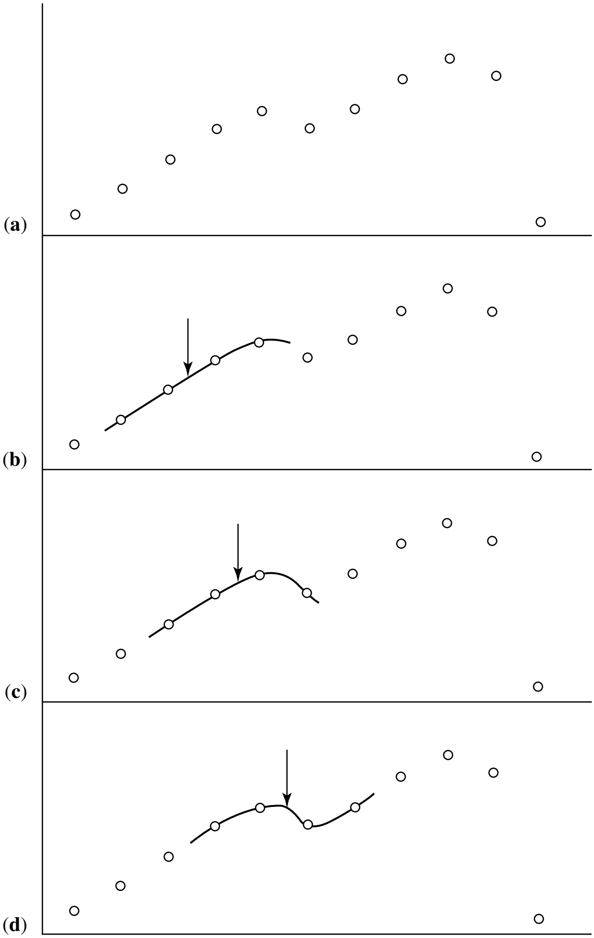


Fig. 3 Convolution with a four-point polynomial spline.

a matrix that will be defined in the following general development of the interpolation method.

A very simple example of interpolation was provided above with the use of a linear function. However, interpolation involving polynomials of higher degree, with more points on either side of the interpolated one is relatively complicated. In effect, the matrix $\boldsymbol{A}$ is then not easily found by inspection.

In the precomputer era a series of experimental points on a graph such as shown in Fig. 3a, was "fitted" with a spline – a sort of flexible ruler that could be adjusted to fit approximately a certain limited number of points. This procedure is illustrated in Fig. 3b–d. The spline forms a smooth curve which can be used by the draftsman to interpolate between successive data points. Furthermore, if there is a certain amount of scatter in the data, a smoothing operation can be carried out, albeit with certain artistic licence.

In computer programs that are now devoted to these problems the interpolation and smoothing of data are special cases of convolution of the data with a set of numerical coefficients, represented for the present by the vector $\boldsymbol{A}_1(x - x_i)$. These coefficients can be determined in advance and placed in memory to be used as needed. If the data points are entered as a vector $\boldsymbol{y}(x_i)$, the convolution can be written in the form

$$\boldsymbol{Y}(x) = \sum_i \boldsymbol{A}_1(x - x_i) \cdot \boldsymbol{y}(x_i). \tag{25}$$

This expression is the discrete form of the convolution integral defined in Eq. (11-13).

In practice the experimental values $\boldsymbol{y}(x_i)$ are usually measured at equally spaced abscissa values and the convolution is applied in succession to limited portions of the experimental data. In principle the equal spacing of data points along the x axis is not necessary, although it is essential in most numerical applications. It is useful to define the difference $\boldsymbol{y} - \boldsymbol{Y} = \boldsymbol{\varepsilon}$, the vector of "errors" at each point. The chosen function $\boldsymbol{Y}(x)$ will be assumed here to be a polynomial of degree $k - 1$, although it can be a more general function. Then, if $\boldsymbol{\theta}$ is a vector composed of the k coefficients in the polynomial

$$\boldsymbol{Y}(x) = \boldsymbol{X}\boldsymbol{\theta}, \tag{26}$$

where $\boldsymbol{X}$ is a $2m \times k$ matrix (with $2m \geq k$) whose elements are powers of x. Specifically this matrix is of the general form

$$\boldsymbol{X} = \begin{pmatrix} (-m+\frac{1}{2})^0 & (-m+\frac{1}{2})^1 & \cdots & (-m+\frac{1}{2})^{k-1} \\ (-m+\frac{3}{2})^0 & & & \\ \vdots & & & \\ (-\frac{1}{2})^0 & & & \vdots \\ (+\frac{1}{2})^0 & & & \\ \vdots & & & \\ (m-\frac{1}{2})^0 & (m-\frac{1}{2})^1 & \cdots & (m-\frac{1}{2})^{k-1} \end{pmatrix}. \tag{27}$$

Then

$$\boldsymbol{\varepsilon} = \boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta} \tag{28}$$

and it is the quantity $S \equiv \tilde{\boldsymbol{\varepsilon}}\boldsymbol{\varepsilon}$ that is minimized in the application of the least-squares criterion used in the previous section. Thus,

$$\frac{\partial S}{\partial \theta_j} = 0, \qquad j = 1, 2, \ldots, k \tag{29}$$

which leads to the matrix relation

$$2\tilde{\boldsymbol{X}}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta}) = \boldsymbol{0}. \tag{30}$$

Its solution for $\boldsymbol{\theta}$ is in the form (problem 8).

$$\boldsymbol{\theta} = (\tilde{\boldsymbol{X}}\boldsymbol{X})^{-1}\tilde{\boldsymbol{X}}\boldsymbol{y} = \boldsymbol{A}\boldsymbol{y} \tag{31}$$

and thus the matrix $\boldsymbol{A}$ [see Eq. (25)] can be constructed from the matrix $\boldsymbol{X}$. It should be noted that the derivatives of order n can also be evaluated, as

$$\left(\frac{\partial^n \boldsymbol{Y}}{\partial \boldsymbol{X}^n}\right)_{X=0} = n!\,\theta_{n+1}\ . \tag{32}$$

The first row of the matrix $\boldsymbol{A}$ consists of the coefficients for the interpolation of the values of y_i, while subsequent rows provide the values of the corresponding derivative coefficients.

The application of the general method can be illustrated by the example shown in Fig. 3. The series of data points is fitted by a polynomial of second degree. Two points will be employed on either side of the point to be interpolated. Thus, $m = 2$ and the matrix $\boldsymbol{X}$ is of the form

$$\boldsymbol{X} = \begin{pmatrix} (-\frac{3}{2})^0 & (-\frac{3}{2})^1 & (-\frac{3}{2})^2 \\ (-\frac{1}{2})^0 & (-\frac{1}{2})^1 & (-\frac{1}{2})^2 \\ (+\frac{1}{2})^0 & (+\frac{1}{2})^1 & (+\frac{1}{2})^2 \\ (+\frac{3}{2})^0 & (+\frac{3}{2})^1 & (+\frac{3}{2})^2 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{3}{2} & \frac{9}{4} \\ 1 & -\frac{1}{2} & \frac{1}{4} \\ 1 & +\frac{1}{2} & \frac{1}{4} \\ 1 & +\frac{3}{2} & \frac{9}{4} \end{pmatrix} \tag{33}$$

and

$$\boldsymbol{A} = (\tilde{\boldsymbol{X}}\boldsymbol{X})^{-1}\tilde{\boldsymbol{X}} = \begin{pmatrix} -\frac{1}{16} & \frac{9}{16} & \frac{9}{16} & -\frac{1}{16} \\ -\frac{3}{10} & -\frac{1}{10} & \frac{1}{10} & \frac{3}{10} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{pmatrix} \tag{34}$$

(problem 9).

The coefficients appearing in the first row of the matrix $\boldsymbol{A}$ provide the weight attributed to each of the four data points that determine the interpolated point indicated by the arrow in Fig. 3b. The elements in the second and third rows are similarly employed if there is an interest in calculating the first and second derivatives, respectively. The determination of the interpolated points at $i + 1$ and $i + 2$ is carried out in the same way, as indicated in Figs. 3c and 3d. This procedure is continued to complete the convolution represented by Eq. (25). The matrices $\boldsymbol{A}$ that have been calculated for polynomials of various degrees and number of points, have been published and are available in certain computer programs.

The interpolation method outlined above can be applied as well to the "smoothing" of experimental data. In this case a given experimental point is replaced by a point whose position is calculated from the values of m points on each side. The matrix $\boldsymbol{X}$ then contains an odd number of columns, namely $2m + 1$. The matrices $\boldsymbol{A}$ have also been tabulated for this application. This smoothing method has been used for a number of years by molecular spectroscopists, who generally refer to it as the method of Savitzky and Golay.*

13.4 THE FOURIER TRANSFORM

13.4.1 The discrete Fourier transform (DFT)

The Fourier transform was defined by Eq. (11-2) as

$$F(k) = \int_{-\infty}^{\infty} f(x) e^{2\pi ikx}\, \mathrm{d}x. \tag{35}$$

As experimental data represented by $f(x)$ are usually symmetrical (even), or can be made so, it is the cosine transform that is appropriate, *viz.*

$$F(k) = \int_{-\infty}^{\infty} f(x)\, cos(2\pi kx)\, \mathrm{d}x. \tag{36}$$

In spectroscopy, for example, the Fourier transform of an interferogram, $f(x)$ is sampled at regular intervals, Δx. Equation (36) is then replaced by the summation

$$F(k) = \Delta x \sum_{m=-M/2}^{m=M/2-1} f(m\Delta x)\, cos(2\pi km\Delta x), \tag{37}$$

where M is the number of points sampled. As $f(x)$ has been assumed to be an even function, Eq. (37) can be written as

---

*George Boris Savitzky, American physical chemist (1925–); Marcel J. E. Golay, Swiss-American physicist (1902–).

$$F(k) = \Delta x \left\{ F(0) + \sum_{m=1}^{m=M/2-1} f(m\Delta x)\, cos(2\pi km\Delta x) \right.$$
$$\left. + f[(M/2)\Delta x]\, cos[2\pi k(M/2)\Delta x] \right\} \tag{38}$$

if an even number of points has been chosen. The total number of terms to be evaluated is then reduced from M to $(M/2) + 1$.

The computer evaluates the cosine functions appearing in Eq. (38) from their series expansions, as given by Eq. (1-34). As M is usually a large number, the time required for the evaluation of the sum can be considerable. However, the arguments of the cosines are simply related because the data points are separated by the constant interval Δx. Given the relation

$$cos\,\alpha + cos\,\beta = 2\,cos\left[\tfrac{1}{2}(\alpha+\beta)\right] cos\left[\tfrac{1}{2}(\alpha-\beta)\right], \tag{39}$$

(problem 10), its application to the present problem can be written as

$$cos[(n+1)\eta] = 2\,cos\,n\eta\,cos\,\eta - cos[(n-1)\eta] \tag{40}$$

where $\eta = \frac{1}{2}(\alpha - \beta)$. This recurrence relation is known as that of Chebyshev. If $\eta = 2\pi km\Delta x$ is the argument of the cosines in the summation [Eq. (38)], all of the other cosines can be calculated from this expression. The result is a considerable saving in the calculation time.

Furthermore, as the output of this calculation is normally in the form of regularly spaced data points, it can be expressed as

$$F(n\Delta k) = \Delta x \sum_{m=1}^{m=M/2-1} f(m\Delta x)\, cos(2\pi mn\Delta k\Delta x). \tag{41}$$

In this case the cosines can be arranged in matrix form, *viz.*

$$\boldsymbol{C} = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots \\ 1 & cos\,u & cos\,2u & cos\,3u & \\ 1 & cos\,2u & cos\,4u & cos\,6u & \\ 1 & cos\,3u & cos\,6u & cos\,9u & \\ \vdots & & & & \ddots \end{pmatrix}, \tag{42}$$

where $u = 2\pi\Delta k\Delta x$. If there are the same number of points in x-space as in k-space, as is usually the case, the matrix $\boldsymbol{C}$ is square and symmetrical with respect to the principal diagonal. With the aid of the Chebyshev recurrence relation the elements of this matrix can be rapidly calculated, once $cos\,u$ has been evaluated. If the ordinate values of $f(x)$ are arranged in the form of a column vector $\boldsymbol{f}$ and those of $(\Delta x)F(k)$ as a vector $\boldsymbol{F}$, the discrete Fourier transform is then calculated by the simple matrix multiplication,

$$\boldsymbol{F} = \boldsymbol{C}\boldsymbol{f}. \tag{43}$$

This numerical method of computing the DFT is sometimes referred to as the slow Fourier transform – by comparison with the fast Fourier transform (FFT) described in the following section.

13.4.2 The fast Fourier transform (FFT)

The fast Fourier transform can be carried out by rearranging the various terms in the summations involved in the discrete Fourier transform. It is, in effect, a special book-keeping scheme that results in a very important simplification of the numerical evaluation of a Fourier transform. It was introduced into the scientific community in the mid-sixties and has resulted in what is probably one of the few significant advances in numerical methods of analysis since the invention of the digital computer.

The basic argument in the FFT algorithm is determined by the initial requirement that

$$k_{max} = \frac{M}{2}\Delta k = \frac{1}{2\Delta x}, \tag{44}$$

where M is the number of measured points. Thus, the point-by-point accumulation of data in k space is made symmetrically with respect to the maximum at k_{max}. Equation (44) corresponds to $\Delta k\Delta x = M^{-1}$. Furthermore, the matrix $\boldsymbol{C}$ is always taken to be square, *viz.* $N = M$ and of rank 2^ℓ, where ℓ is an integer. Under these conditions the cosines appearing in the matrix $\boldsymbol{C}$ will all be of the form $cos(2\pi mn/M)$. Here, the (independent) indices n and m have been chosen to run from zero to $M - 1$. In this case the general expression for the discrete Fourier transform [Eq. (41)] can be written as

$$F(n) = \Delta x \sum_{m=0}^{M-1} f(m)\, cos\left(\frac{2\pi mn}{M}\right). \tag{45}$$

Furthermore, the choice of the cosine transform implies that $f(m)$ is symmetrical about its maximum value; thus, $f(m - M) = f(m)$.

With the arguments of the preceding paragraph in mind it becomes possible to construct the functions $F(n)$ in k space. This procedure is best explained with the aid of an example. Consider the simple case in which $M = 8 = 2^3$. With $\Delta x = \frac{1}{2}$, Eq. (45) yields the expressions

$$\begin{aligned} F(0) &= \tfrac{1}{2}\{f(0) + f(4) + f(2) + f(6) + f(1) + f(5) + f(3) + f(7)\} \\ &= \tfrac{1}{2}\{f(0) + f(4) + 2f(2) + 2f(1) + 2f(3)\}, \end{aligned} \tag{46}$$

$$\begin{aligned} F(1) &= \tfrac{1}{2}\{f(0) - f(4) + 0f(2) + 0f(6) \\ &\qquad + [f(1) - f(5) - f(3) + f(7)]\, cos\, \tfrac{\pi}{4}\} \\ &= \tfrac{1}{2}\{f(0) - f(4) + 2[f(1) + -f(5)]\, cos\, \tfrac{\pi}{4}\}, \end{aligned} \tag{47}$$

$$F(2) = \tfrac{1}{2}\{f(0) + f(4) - f(2) - f(6)\}, \tag{48}$$

$$F(3) = \tfrac{1}{2}\{f(0) - f(4) - [f(1) - f(5) - f(3) + f(7)]\, cos\, \tfrac{\pi}{4}\} \tag{49}$$

and

$$F(4) = \tfrac{1}{2}\,\{f(0) + f(4) + f(2) + f(6) - f(1) - f(5) - f(3) - f(7)\} \tag{50}$$

(problems 11).

The order in which the functions $f(m)$ are presented in the above relations is specific. First, note that all of the functions of even values of m are specified before those of odd values. Moreover, the order employed here is referred to as reverse binary order,* which does not correspond to the order that might be intuitively established, namely, $m = 0, 1, 2, \ldots, 7$. Furthermore, each is multiplied by a value of $cos(2\pi nm/8)$, as $M = 8$ in this case. Clearly, Eqs. (46–50) can be recast in matrix form. However, with the addition of the symmetry conditions $F(5) = F(3)$, $F(6) = F(2)$ and $F(6) = F(1)$ the appropriate 8×8 matrix $\boldsymbol{C}$ can be easily constructed. On the other hand, if the inverse binary order is also imposed on the elements of the vector $F(n)$, a considerable simplification results.

Continuing with the eight-point transform, Eq. (43) can be written in the form

$$\begin{pmatrix} F(0) \\ F(4) \\ F(2) \\ F(6) \\ F(1) \\ F(5) \\ F(3) \\ F(7) \end{pmatrix} = \frac{1}{2} \left(\begin{array}{cccc:cccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ \hdashline 1 & -1 & 0 & 0 & c & -c & -c & c \\ 1 & -1 & 0 & 0 & -c & c & c & -c \\ 1 & -1 & 0 & 0 & -c & c & c & -c \\ 1 & -1 & 0 & 0 & c & -c & -c & c \end{array}\right) \begin{pmatrix} f(0) \\ f(4) \\ f(2) \\ f(6) \\ f(1) \\ f(5) \\ f(3) \\ f(7) \end{pmatrix}, \tag{51}$$

---

*In binary algebra [Boolean, after George Boole, British mathematician (1815–1864)] the indices $0, 1, 2, \ldots, 7$ are represented by 000, 001, 010, 011, 100, 101, 110, and 111, respectively. The reversal of these binary numbers yields the values of m in the order indicated in Eqs. (46)–(50).

(problem 12). Note that the matrix $\boldsymbol{C}$ is now symmetric with respect to the principal diagonal. Furthermore, there is symmetry with respect to the parity of both m and n. In the 4×4 block that corresponds to both m and n even the columns occur in identical pairs, while in the block with both m and n odd the result is analogous, although the signs are reversed. It is apparent that only four matrix elements need be evaluated, *viz.* $c = cos(\pi/4)$, plus the trivial ones $cos\, 0 = 1$, $cos\, \pi/2 = 0$ and $cos\, \pi = -1$. The matrix $\boldsymbol{C}$ given in Eq. (51) should be compared with that obtained by application of Eq. (42). Note that the factor $\frac{1}{2}$ is just $\Delta x = 1/\ell$.

It is instructive to consider a specific example of the method outline above. The triangle function $(1/\ell) \wedge (x/\ell)$ was discussed in Section 11.1.2. It was pointed out there that it arises in dispersive spectroscopy as the slit function for a monochromator, while in Fourier-transform spectroscopy it is often used as an apodizing function.* Its Fourier transform is the function $sinc^2$, as shown in Fig. (11-2). The eight points employed to construct the normalized triangle function define the matrix

$$\boldsymbol{f} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{3}{8} \\ \frac{1}{8} \\ \frac{1}{8} \\ \frac{3}{8} \end{pmatrix} \tag{52}$$

where it is essential to preserve the order of the elements as given in Eq. (51). Multiplication of the vector $\boldsymbol{f}$ by the matrix $\boldsymbol{C}$ of Eq. (51) yields

$$\boldsymbol{F} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \frac{1}{4}(1+c) \\ \frac{1}{4}(1-c) \\ \frac{1}{4}(1-c) \\ \frac{1}{4}(1+c) \end{pmatrix}, \tag{53}$$

---

*An apodizing function is employed to reduce oscillations in an observed spectrum due to discontinuities at the ends of an interferogram.

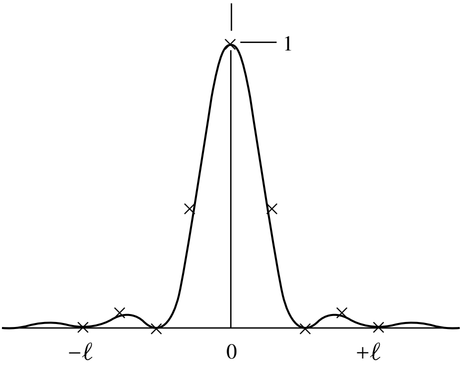


Fig. 4 The Fourier transform of the triangle, $(1/\ell)\bigwedge(x/\ell)$. The points calculated with the use of Eq. (53) are indicated by ×.

which is the desired Fourier transform. This result is compared with the function *sinc*^{2} in Fig. 4. Considering the limited number of points employed, this agreement is excellent. Clearly, it would be necessary to carry out the transform with a larger number of points to obtain a more convincing description of the function *sinc*^{2}.

The symmetry and simplicity of the matrix $\boldsymbol{C}$ (and hence the extreme rapidity of the FFT) is determined by the particular order employed in both the input vector $\boldsymbol{f}$ and the output $\boldsymbol{F}$. Thus, both sets of data must be rearranged from what would be normally expected. While this problem represents an inconvenience for a programmer, it is carried out automatically in available programs. Although it would probably go un-noticed by the user, it is important for him or her to understand the fundamental algorithm of the FFT, which is based on the inverse binary order explained here.

13.4.3 An application: interpolation and smoothing

Both interpolation and smoothing of experimental data are of particular importance in all branches of spectroscopy. One approach to this problem was considered in Section 13.3. However, with the development of the FFT another, often more convenient, method has become feasible. The basic argument is illustrated in Fig. 5. Given a particular problem whose solution may appear to be difficult, it is sometimes possible to resolve it *via* recourse to the Fourier transform.

Consider the problem of smoothing an experimental curve, such as represented in Fig. 6a. It might well correspond to a spectrum, as observed in absorption, emission or, say, Raman scattering. The noise is, usually at least,

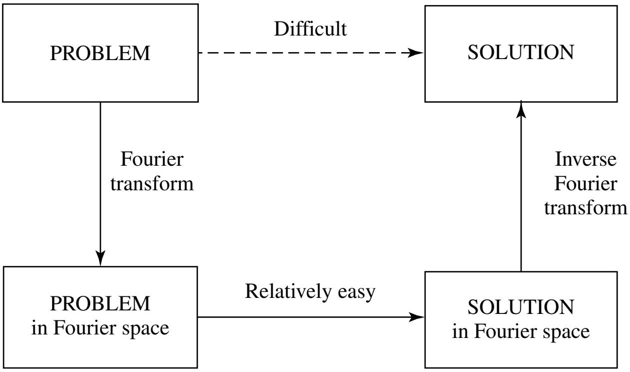


Fig. 5 The solution of a problem with the aid of the Fourier transform.

largely random. As it can be reasonably represented by a normal distribution, it would seem appropriate to smooth the observed data by convolution with a Gaussian function. This operation is conveniently carried out by first transforming the data into Fourier space, where they can simply be multiplied by the appropriate Gaussian, as the latter is of course the Fourier transform of the original Gaussian (see Fig. 11-3). The product is subsequently transformed back as a smoothed spectrum. An example is shown in Fig. 6. If convoluted by a Gaussian of width 5 points, the experimental spectrum of Fig. 6a is smoothed as in Fig. 6b. Further smoothing, for example by a Gaussian of width 30 points, results in a substantial loss of information. Thus, as in the case of polynomial smoothing, this method must be used with discretion.

The principle presented above can also be applied to interpolate points in an experimental profile. If the original function $\boldsymbol{F}$ (a spectrum, for example) is transformed with the use of the FFT algorithm, the result is a function of the same number of points in Fourier space. It might be, for example, the original interferogram $\boldsymbol{f}$ that was used to generate the spectrum. The number of points can be augmented by simply adding zeros to the vector $\boldsymbol{f}$. If the number of points is doubled, the result of carrying out the inverse transform is to produce the vector $\boldsymbol{F}$ with twice as many points as before. This procedure corresponds exactly to midpoint interpolation of the original spectrum by the function *sinc*. This result should become evident if it is recalled that the *sinc* function is the Fourier transform of the boxcar (Fig. 11-1), whose width has been doubled by the operation of “zero filling”. Obviously, no new information is obtained by this procedure, but the result may be of esthetic value in the presentation of the spectrum. This method yields better results than the more usual polynomial interpolation method presented in Section 13.3.

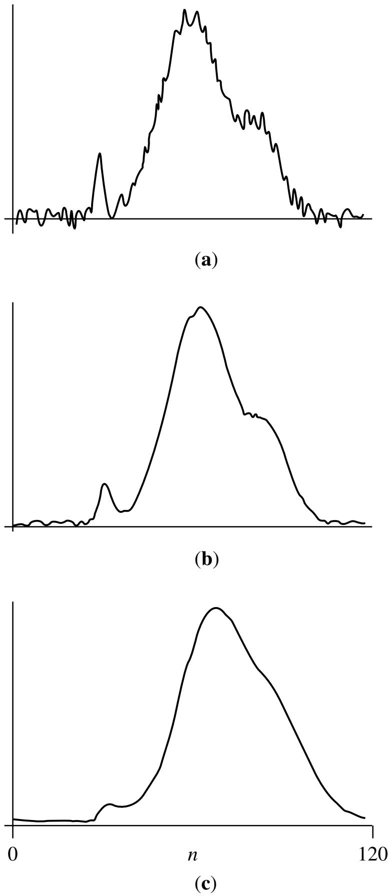


Fig. 6 Gaussian convolution of an experimental profile: (**a**) raw spectrum; (**b**) after convolution by a Gaussian of width 5 points; (**c**) after convolution by a Gaussian of 30 points. The ordinate scale is arbitrary.

13.5 NUMERICAL INTEGRATION

The numerical evaluation of definite integrals can be carried out in several ways. However, in all cases it must be assumed that the function, as represented by a table of numerical values, or perhaps a known function, is well behaved. While this criterion is not specific, it suggests that the functions having pathological problems, *e.g.* singularities, discontinuities, . . ., may not survive under the treatment in question.

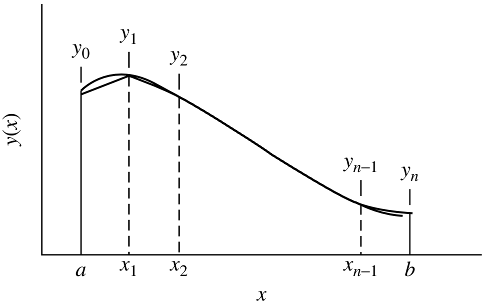


Fig. 7 The trapezoids.

13.5.1 The trapezoid rule

Consider a function $y(x)$ as shown in Fig. 7. If the interval of integration, say from a to b, is divided in n equal intervals,

$$x_k = a + k\Delta x, \tag{54}$$

where $k = 0, 1, 2, \ldots, n$, $\Delta x = (b-a)/n$ and $y_k = y(x_k)$. If $y(x)$ is expanded in a Taylor series (Section 2.9),

$$y(x) = y(x_0) + (x - x_0)y'(x_0) + \tfrac{1}{2!}(x - x_0)^2 y''(x_0) + \cdots. \tag{55}$$

The integral of this expression from x_0 to x_1 is given by

$$\begin{aligned}\int_{x_0}^{x_1} y(x)\,\mathrm{d}x &= (x - x_0)y(x_0) + \left[\frac{(x-x_0)^2}{2!}y'(x_0)\right]_{x_0}^{x_1} \\ &\quad + \left[\frac{(x-x_0)^3}{3!}y''(x_0)\right]_{x_0}^{x_1} + \cdots \\ &= \Delta x\, y(x_0) + \left[\frac{\overline{\Delta x}^2}{2!}y'(x_0)\right] + \left[\frac{\overline{\Delta x}^3}{3!}y''(x_0)\right] + \cdots.\end{aligned} \tag{56}$$

From Eq. (55)

$$y_1 = y(x_1) = y_0 + \Delta x\, y'(x_0) + \frac{\overline{\Delta x}^2}{2!}y''(x_0) + \cdots\ , \tag{57}$$

which when multiplied by $\overline{\Delta x}^2/2$ and substituted in Eq. (56) yields

$$\int_{x_0}^{x_1} y(x)\,\mathrm{d}x = \frac{\Delta x}{2}[y(x_0) + y(x_1)] - \frac{\overline{\Delta x}^3}{12}y''(x_0) + \cdots. \tag{58}$$

The integral over the entire region can then be written as

$$\int_a^b y(x)\,\mathrm{d}x = \int_{x_0}^{x_n} y(x)\,\mathrm{d}x \approx \frac{\Delta x}{2}[y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n] + \mathcal{O}(\overline{\Delta x}^2), \tag{59}$$

where the term $\mathcal{O}(\overline{\Delta x}^2) = -(\overline{\Delta x}^2/12)[y'(b) - y'(a)]$ represents the error in the simple trapezoid method of numerical integration.

13.5.2 Simpson's rule*

In the method presented in the previous section each vertical "slice" was defined by two successive points, $x_0, x_1; x_1, x_2$; *etc.* If now the successive points are selected three-by-three, they can be connected by a parabola. The approximate integral over the first two slices can then be written as

$$\int_{x_0}^{x_2} y(x)\,\mathrm{d}x \approx \frac{\Delta x}{3}[y(x_0) + 4y(x_1) + y(x_2)]. \tag{60}$$

The correction to this expression involves multiple derivatives, although the basic equation, Eq. (60), does not. The development of this result, as above for the trapezoid rule, leads to the relation for the integral over the range a to b in the form

$$\int_a^b y(x)\,\mathrm{d}x \approx \frac{\Delta x}{3}[y(x_0) + 4y(x_1) + 2y(x_2) + 4y(x_3) + \cdots + 2y(x_{n-2}) + 4y(x_{n-1}) + y(x_n)]. \tag{61}$$

It should be noted that n has been assumed here to be even. Equation (61), without the inclusion of the correction term in $\overline{\Delta x}^4$, is the one usually used in the numerical evaluation of integrals. When higher precision is required, and a suitable computer is available, the algorithm described in the following section can be employed.

13.5.3 The method of Romberg†

The two well-known methods of numerical integration described in the previous sections can be generalized. Represent the sum on the right-hand side of Eq. (59) as $S_0(n)$. This function converges but very slowly towards

---

*Thomas Simpson, English mathematician (1710–1761).

†Werner Romberg, German mathematician (1909–).

the exact value of the integral as $n \to \infty$. However, the following method is much more efficient.

Reconsider Eq. (59) in the form

$$S_0(n) = \int_a^b f(x)\,\mathrm{d}x + C(\overline{\Delta x}^2) + \mathcal{O}(\overline{\Delta x}^4), \tag{62}$$

where C is a constant. If the number of intervals n is now doubled, this expression becomes

$$S_0(2n) = \int_a^b f(x)\,\mathrm{d}x + \tfrac{1}{4}C(\overline{\Delta x}^2) + \mathcal{O}(\overline{\Delta x}^4). \tag{63}$$

By eliminating the constant C between Eqs. (62) and (63) the relation

$$\begin{aligned}\int_a^b f(x)\,\mathrm{d}x &= \frac{4S_0(2n) - S_0(n)}{3} + \mathcal{O}(\overline{\Delta x}^4)\\ &= S_1(2n) + \mathcal{O}(\overline{\Delta x}^4)\end{aligned} \tag{64}$$

can easily be established. If the correction terms are neglected, this result is equivalent to Simpson's rule for the division of the interval a,b in $2n$ equal slices. If the process of halving the intervals is continued, the expression

$$S_0(n) = \frac{\overline{\Delta x}}{2}[y(x_0) + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n] \qquad n = 1, 2, 4, 8, \ldots \tag{65}$$

can be obtained for the application of the simple trapezoid rule for each value of n. This result is the starting point for the application of Romberg's method. It is continued by application of the recursion relation that is obtained by generalizing of Eq. (64). It is given by

$$S_m(2n) = \frac{4^m S_{m-1}(2n) - S_{m-1}(n)}{4^m - 1}, \tag{66}$$

with $m = 1, 2, 3, \ldots$ and $n = 2^{m-1}, 2^m, 2^{m+1}, \ldots$.

As an example of the application of Romberg's method, consider the integral

$$I(T/\theta_D) = \int_0^{T/\theta_D} \frac{x^3}{e^x - 1}\,\mathrm{d}x, \tag{67}$$

that arises in Debye's theory of the heat capacity of solids.* In Eq. (67), T is the absolute temperature and θ_D is referred to as the Debye temperature. In the low-temperature limit the integral in Eq. (67) is given approximately by

---

*Petrus Debye, Dutch-American physicist and chemist (1884–1966).

$\frac{1}{3}(T/\theta_D)^3$. This limiting expression is known as Debye's third-power law for the heat capacity (problem 15). It is employed in thermodynamics to evaluate the low-temperature contribution to the absolute entropy.

The integral in Eq. (67) cannot be evaluated analytically. However, for a given upper limit T/θ_D, it can be calculated, in principle to any desired precision, with the application of the methods outlined above. The results for $T/\theta_D = 1.6$ are summarized in Table 1.

Table 1 Evaluation of Eq. (67) with $T/\theta_D = 1.6$.

n	(Trapezoid) $S_0(n)$	(Simpson) $S_1(n)$	(Milne) $S_2(n)$	$S_3(n)$
1	0.8289332462284			
2	0.74868638720142	0.72193743419243		
4	0.724310280204	0.71618491120487	0.7158014096724	
8	0.71795245734086	0.71583318305313	0.71580973450967	0.71580986664995
16	0.71634660087305	0.71581131538377	0.71580985753913	0.71580985949197
32	0.71594411304265	0.7158099504325	0.71580985943573	0.71580985946584
64	0.71584342712365	0.71580986515063	0.71580985946513	0.71580985946559

The slowest part of the construction of this table is the evaluation of the entries in the first column. The simple trapezoid rule, as given by Eq. (65), is applied with successive sectioning of the slices. It can be seen that by descending the column a limiting value can, in principle, be obtained. However, the convergence is very slow. With the use of the recursion relation given by Eq. (66), each successive pair of entries in the first column can be employed to calculate the values presented in the second column. The results shown for this example are equal to those obtained by Simpson's method [Eq. (61)].

The third column of Table 1 is calculated by applying the recursion relation to the values shown in the second column, *etc.* It corresponds to the method of Milne.* It is apparent that the convergence becomes much more rapid with each successive column. For this particular example the same limiting values is obtained as either n or m becomes very large.

13.6 ZEROS OF FUNCTIONS

13.6.1 Newton's method

Given a function $f(x)$, if its derivatives can be evaluated numerically, Newton's method can often serve as an algorithm for the determination of

---

*William. E. Milne, American mathematician (1890–1971).

the zeros of the equation $f(x) = 0$. Assume that x_0 is an estimated value of one of the roots. Then, at least in principle, an improved value of the root is given by

$$x = x_0 + \Delta x, \tag{68}$$

where $\Delta x = -f(x_0)/f'(x_0)$. The next approximation is found by replacing x_0 by x in Eq. (68) to get a new value of Δx. This procedure is continued as long as is necessary to obtain the desired accuracy. Usually, after a few successive approximations, the value of the derivative will change little; hence, $f'(x)$ need not be recalculated each time. It should be obvious that the solution will be found more rapidly if the initial value x_0 is wisely chosen.

13.6.2 The bisection method

In the application of the bisection method it is assumed only that the function $f(x)$ is continuous. It requires that two initial values of x, say x_a and x_b, be chosen so that they straddle the desired zero. Thus, $f(x_\mathrm{a})$ and $f(x_\mathrm{b})$ will have opposite signs and their product will be negative. Now, take the midpoint $x_\mathrm{m} = (x_\mathrm{a} + x_\mathrm{b})/2$ and calculate $f(x_\mathrm{m})$. If, for example, the product $f(x_\mathrm{a})f(x_\mathrm{m}) < 0$, the desired root lies between x_a and x_m. The midpoint between these two limits is then calculated and the process is repeated to the desired degree of accuracy. Here again, the better the choice of the initial limits, the fewer the number of bisections that will be required.

13.6.3 The roots: an example

The function $f(x) = (5-x)e^x - 5$ arises in the theory of black-body radiation. Obviously, it has a zero at $x = 0$. A plot of this function (Fig. 8) shows that it has a second zero near $x = 5$. As this function appears to be well behaved in this region, Newton's method might be expected to yield a value for the second root.

If, as a guess, the initial value of x is chosen to be $x_0 = 4.5$, convergence to the value $x = 4.96511$ will occur within a few iterations. On the other side, where $x > 5$, even wilder guesses will yield the same, correct answer. However, if $x_0 = 4$ is taken as a starting point, disaster will result. Reference to the plot of this function in Fig. 8 indicates that this point is at the maximum. As the slope is then equal to zero, the computer will yield a "division by zero" message for the calculation of Δx and the method fails. Of course if $x_0 = 3$ were chosen as the initial value, the procedure will converge to the root at $x = 0$. Clearly, the function must be plotted if such pitfalls are to be avoided.

As the bisection method does not depend on the derivatives of the function in question, it can be applied with confidence, even if there are stationary points within the chosen limits, x_a and x_b. However, convergence is often

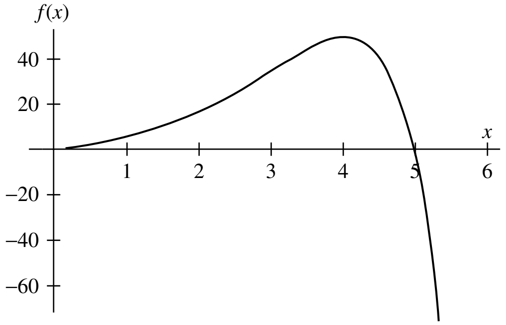


Fig. 8 The function $f(x) = (5-x)e^x - 5$ as a function of x.

somewhat slower. It is to be emphasized that it is assumed in this method that the function is continuous between the chosen limits. Here again, it is essential to plot the function before undertaking the evaluation of its roots.

A final remark should be added that applies to both of the methods outlined above. As both are iterative, any computer program must specify either the number of iterations or the precision of the desired result. Or better, both should be included and employed – whichever comes first.

PROBLEMS

1. Make the indicated substitution to yield Eq. (2).
2. Develop the logarithms in Eqs. (2) and (3) to obtain Eq. (4).
3. Show that the Gaussian function given by Eq. (4) is correctly normalized.
4. Verify Eqs. (12) and (13).
5. Derive the expressions for m and b in Eq. (16). Ans. Eqs. (20) and (21)
6. Verify the least-squares fit to the data given in Fig. 1-1.
7. Show that in the application of linear, midpoint interpolation

$$\boldsymbol{X} = \begin{pmatrix} \left(-\frac{1}{2}\right)^0 & \left(-\frac{1}{2}\right)^1 \\ \left(+\frac{1}{2}\right)^0 & \left(+\frac{1}{2}\right)^1 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} \\ 1 & +\frac{1}{2} \end{pmatrix}$$

and thus $\boldsymbol{A}$ is given by Eq. (24).

8. Derive the general expression for $\boldsymbol{\theta}$ [Eq. (31)].

9. Verify the matrix $\boldsymbol{A}$ given by Eq. (34).

10. Derive Eq. (39).

11. Check the relations given by Eqs. (46)–(50).

12. Construct the matrix given in Eq. (51).

13. Carry out the matrix multiplication indicated to obtain Eq. (53).

14. Prove Eq. (64).

15. Derive Debye's third-power law.

Appendix I: The Greek Alphabet

Greek letter	Greek name	Approximate English equivalent
Α α	Alpha	a
Β β	Beta	b
Γ γ	Gamma	g
Δ δ	Delta	d
Ε ε	Epsilon	ĕ
Ζ ζ	Zeta	z
Η η	Eta	ē
Θ θ, ϑ	Theta	th
Ι ι	Iota	i
Κ κ	Kappa	k
Λ λ	Lambda	l
Μ μ	Mu	m
Ν ν	Nu	n
Ξ ξ	Xi	x
Ο ο	Omicron	ŏ
Π π	Pi	p
Ρ ρ	Rho	r
Σ σ, ς	Sigma	s
Τ τ	Tau	t
Υ υ	Upsilon	u
Φ ϕ, φ	Phi	ph
Χ χ	Chi	ch
Ψ ψ	Psi	ps
Ω ω	Omega	ō

Appendix II: Dimensions and Units

The numerical values of most physical quantities are expressed in terms of units. The distance between two points, for example, can be specified by the number of meters (or feet, Ångströms, *etc.*). Similarly, time can be expressed in seconds, days or, say, years. However, the number of days per year varies from one year to another. The quantities, distance (length) and time, as well as mass, are usually chosen to be primary quantities. In terms of them Newton's second law for the force on an object, can be written as force = mass.distance/(time)^2. The definition of the primary quantities allows dimensional expressions to be written, such as [force] = MLT^{-2} in the present example. Note, however, that in everyday life one speaks of the weight of an object (or a person). Of course the weight is not the mass, but rather the force acting on the object by the acceleration due to gravity: [acceleration] = LT^{-2}.

The dimensional expressions given above are determined by the particular choice of primary quantities. In the international system (SI) the base units of mass, length and time have been chosen to be kilogram, meter and second, respectively. Then, the newton (N, a derived unit) is the unit of force. In the cgs system the centimeter, gram and second are considered to be the base units, leading to the force expressed in dynes. In this particular example the dimensional equation [force] = MLT^{-2} applies to either choice of units. As a force that produces a change in distance involves work or energy, their dimensions are given by $\mathrm{ML}^2\mathrm{T}^{-2}$. The unit of energy in SI is the joule (J), while in cgs it is the erg (note that 1 erg = 10^{-7} J).

The primary quantities M, L, T are sufficient to describe most problems in mechanics. In thermodynamics and other thermal applications it is customary to add an absolute temperature. In this case the dimension of the Boltzmann constant, for example, is given by $[k] = \mathrm{ML}^2\mathrm{T}^{-2}\theta^{-1}$, where the symbol θ is used here for the dimension of the absolute or thermodynamic temperature.

The dimensions of units in electricity and magnetism are the origin of much confusion. In the days when mechanical and thermal quantities were expressed in cgs, two different systems were introduced for the electrical and magnetic quantities. They are the esu (electrostatic units) and the emu (electromagnetic units), respectively. Their addition to the cgs system results in a hybrid that is usually referred to as the Gaussian system. An apparent advantage of the Gaussian system is the disappearance of the factor $4\pi\varepsilon_0$ which is forever present in SI in problems involving inherent spherical symmetry. On the other hand, in the Gaussian system a given quantity usually has different values in esu and emu. It then becomes necessary to introduce various powers of the velocity of light (c) to assure internal consistency.

The so-called atomic units are often employed in quantum mechanical calculations. They are combinations of fundamental constants that are treated as if they were units. The base dimensions are chosen to be mass, length, charge and action. They are respectively the rest mass of the electron (m_e), the radius of the first Bohr orbit ($a_0 = 4\pi\varepsilon_0\hbar^2/m_e e^2$), the elementary charge (e) and the action ($\hbar = h/2\pi$, where h is Planck's constant). The corresponding energy, given by $E_h = \hbar^2/m_e a_0^2 = m_e e^4/(4\pi\varepsilon_0)^2\hbar^2$, is expressed in hartree.

The following tables summarize the units used in this book. For more extensive tabulations, the reader is referred to the "Green Book", Ian Mills, *et al.* (eds), "Quantities, Units and Symbols in Physical Chemistry", Blackwell Scientific Publications, London (1993).

Table 1 The SI base units.

Physical quantity (dimension in SI)	SI unit	Symbol
length (L)	meter	m
mass (M)	kilogram	kg
time (T)	second	s
thermodynamic or absolute temperature (θ)	kelvin	K
electric current (A)	ampere	A
amount of substance (mol)	mole	mol
luminous intensity (cd)	candela	cd

Table 2 Some derived units.

Physical quantity	SI unit	Symbol	Expression in SI	Dimension in SI
frequency	hertz	Hz	s^{-1}	T^{-1}
force	newton	N	$m\ kg\ s^{-2}$	$M\ L\ T^{-2}$
pressure	pascal	Pa	$N\ m^{-2}$	$M\ L^{-1}T^{-2}$
energy, work	joule	J	$N\ m$	$M\ L^2T^{-2}$
power	watt	W	$J\ s^{-1}$	$M\ L^2T^{-3}$
electric charge	coulomb	C	$A\ s$	$A\ T$
electric potential (emf)	volt	V	$J\ C^{-1}$	$M\ L^2T^{-3}A^{-1}$
electric resistance	ohm	Ω	$V\ A^{-1}$	$M\ L^2T^{-3}A^{-2}$
electric capacitance	farad	F	$C\ V^{-1}$	$M^{-1}L^{-2}T^4A^2$

Table 2 (*Continued*).

Physical quantity	SI unit	Symbol	Expression in SI	Dimension in SI
electric field strength			$\mathrm{V\ m^{-1}}$	$\mathrm{L^2T^{-3}A^{-1}}$
Celsius temperature	degree Celsius	C	K	θ
density			$\mathrm{kg\ m^{-3}}$	$\mathrm{M\ L^{-3}}$
molar volume			$\mathrm{m^3mol^{-1}}$	$\mathrm{L^3mol^{-1}}$
wavenumber[a]			$\mathrm{m^{-1}}$	$\mathrm{L^{-1}}$
permittivity (vacuum)		$\varepsilon(\varepsilon_0)$	$\mathrm{F\ m^{-1}}$	$\mathrm{M^{-2}L^{-2}T^4A^2}$

[a] Wavenumber is invariably expressed in $\mathrm{cm^{-1}}$.

Table 3 Prefixes in SI.

Submultiple	Prefix	Symbol	Multiple	Prefix	Symbol
10^{-1}	deci	d	10	deca	da
10^{-2}	centi	c	10^2	hecto	h
10^{-3}	milli	m	10^3	kilo	k
10^{-6}	micro	μ	10^6	mega	M
10^{-9}	nano	n	10^9	giga	G
10^{-12}	pico	p	10^{12}	tera	T
10^{-15}	femto	f	10^{15}	peta	P
10^{-18}	atto	a	10^{18}	exa	E

Table 4 Some of the fundamental constants in the SI system.

Quantity	Symbol	Value
permeability of vacuum	μ_0	$4\pi 10^{-7}\mathrm{H\ m^{-1}}$ or $\mathrm{V\ s\ A^{-1}m^{-1}}$ (defined)
permittivity of vacuum	$\varepsilon_0 = (m_0c^2)^{-1}$	$8.854\ 187\ 816\ldots 10^{-12}\mathrm{F\ m^{-1}}$ or $\mathrm{C^2}J^{-1}\mathrm{m^{-1}}$
speed of light in vacuum	c_0	299 792 458 $\mathrm{m\ s^{-1}}$ (defined)
Planck constant	h	6.626 075 5(40) 10^{-34} J s
elementary charge	e	1.602 177 33(49) 10^{-19} C
electron rest mass	m_e	9.109 389 7(54) 10^{-31} kg
proton rest mass	m_p	1.672 623 1(10) 10^{-27} kg
Avogadro constant	N_A	6.022 136 7(36) $10^{23}\mathrm{mol^{-1}}$
Boltzmann constant	k	1.380 658 (12) $10^{-23}\mathrm{J\ K^{-1}}$
gas constant	R	8.314 510 (70) $\mathrm{J\ K^{-1}mol^{-1}}$
zero of the Celsius scale		273.15 K (defined)
Bohr radius	$a_0 = 4\pi\varepsilon_0\hbar^2/m_ee^2$	5.291 772 49(24) 10^{-11} m
Hartree energy	$E_h = \hbar^2/m_ea_0^2$	4.359 748 2(26) 10^{-18} J
Rydberg constant	$R_\infty = E_h/2hc_0$	1.097 373 153 4(13) $10^7\mathrm{m^{-1}}$

This Page Intentionally Left Blank

Appendix III: Atomic orbitals

The angular functions presented in Table 1 are derived from the wavefunctions for one-electron systems, *e.g.* the hydrogen atom. However, they can be

Table 1 The normalized atomic orbitals.

ℓ	m	symbol[a]	normalizing factor	angular function
0	0	s	$\frac{1}{2}\sqrt{\frac{1}{\pi}}$	
1	0	p_z	$\frac{1}{2}\sqrt{\frac{3}{\pi}}$	$cos\,\theta$
	± 1	p_x	$\frac{1}{2}\sqrt{\frac{3}{\pi}}$	$sin\,\theta\,cos\,\varphi$
		p_y	$\frac{1}{2}\sqrt{\frac{3}{\pi}}$	$sin\,\theta\,sin\,\varphi$
2	0	d_{z^2}	$\frac{1}{4}\sqrt{\frac{5}{\pi}}$	$3\,cos^2\theta - 1$
	± 1	d_{xz}	$\frac{1}{2}\sqrt{\frac{15}{\pi}}$	$sin\,\theta\,cos\,\theta\,cos\,\varphi$
		d_{yz}	$\frac{1}{2}\sqrt{\frac{15}{\pi}}$	$sin\,\theta\,cos\,\theta\,sin\,\varphi$
	± 2	$d_{x^2-y^2}$	$\frac{1}{4}\sqrt{\frac{15}{\pi}}$	$sin^2\theta\,cos\,2\varphi$
		d_{xy}	$\frac{1}{4}\sqrt{\frac{15}{\pi}}$	$sin^2\theta\,cos\,2\varphi$
3	0	f_{z^3}	$\frac{1}{4}\sqrt{\frac{7}{\pi}}$	$5\,cos^3\theta - 3\,cos\,\theta$
	± 1	f_{xz^2}	$\frac{1}{8}\sqrt{\frac{42}{\pi}}$	$sin\,\theta(5\,cos^2\theta - 1)\,cos\,\varphi$
		f_{yz^2}	$\frac{1}{8}\sqrt{\frac{42}{\pi}}$	$sin\,\theta(5\,cos^2\theta - 1)\,sin\,\varphi$
	± 2	$f_{z(x^2-y^2)}$	$\frac{1}{4}\sqrt{\frac{105}{\pi}}$	$sin^2\theta\,cos\,\theta\,cos\,2\varphi$
		f_{zxy}	$\frac{1}{4}\sqrt{\frac{105}{\pi}}$	$sin^2\theta\,cos\,\theta\,sin\,2\varphi$
	± 3	$f_{x(x^2-3y^2)}$	$\frac{1}{8}\sqrt{\frac{70}{\pi}}$	$sin^3\theta\,cos\,3\varphi$
		$f_{y(3x^2-y^2)}$	$\frac{1}{8}\sqrt{\frac{70}{\pi}}$	$sin^3\theta\,sin\,3\varphi$

[a] Note that d_{z^2} is the short notation for $d_{2z^2-x^2-y^2}$, as it appears in the cubic point groups (Appendix VII). Similarly, f_{z^3}, f_{xz^2} and f_{yz^2} are the abbreviated symbols for $f_{z(5z^2-3r^2)}$, $f_{x(5z^2-r^2)}$ and $f_{y(5z^2-r^2)}$, respectively, where $r^2 = x^2 + y^2 + z^2$.

applied as “atomic orbitals” to polyelectron systems, as the appropriate wavefunctions are written as products of those for hydrogen. Although the radial part of the functions is modified by the effect of electron−electron interaction,

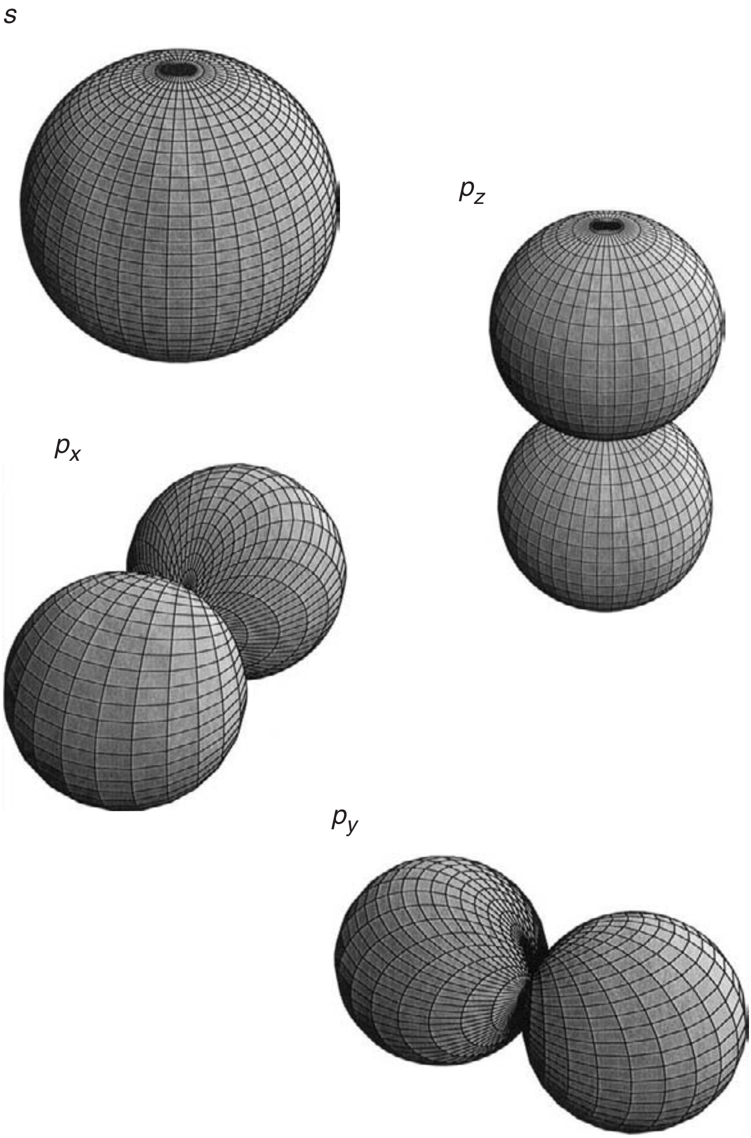


Fig. 1 Some of the atomic orbitals.

the angular parts retain their symmetry. In effect, the functions given here can be employed to describe the basic symmetry of polyelectronic atoms and ions.

Three-dimensional representations of the angular dependence of some of the atomic orbitals are shown in Fig. 1. The coordinate axes and angles are defined in Fig. 6-5.

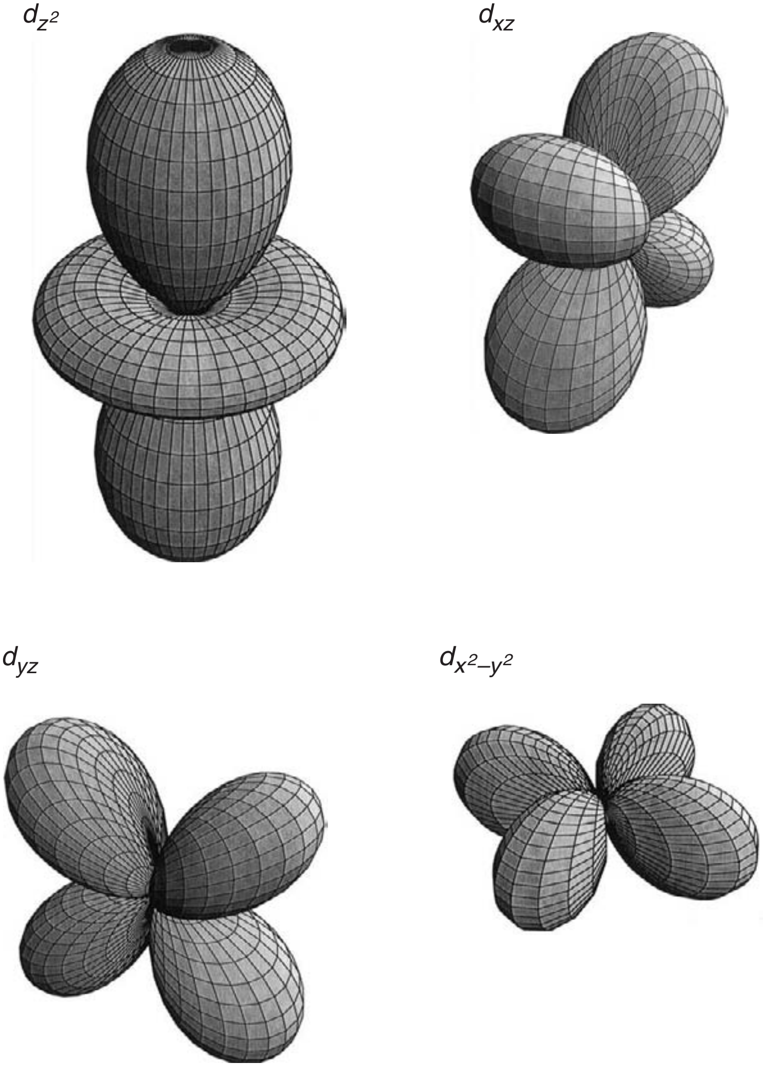
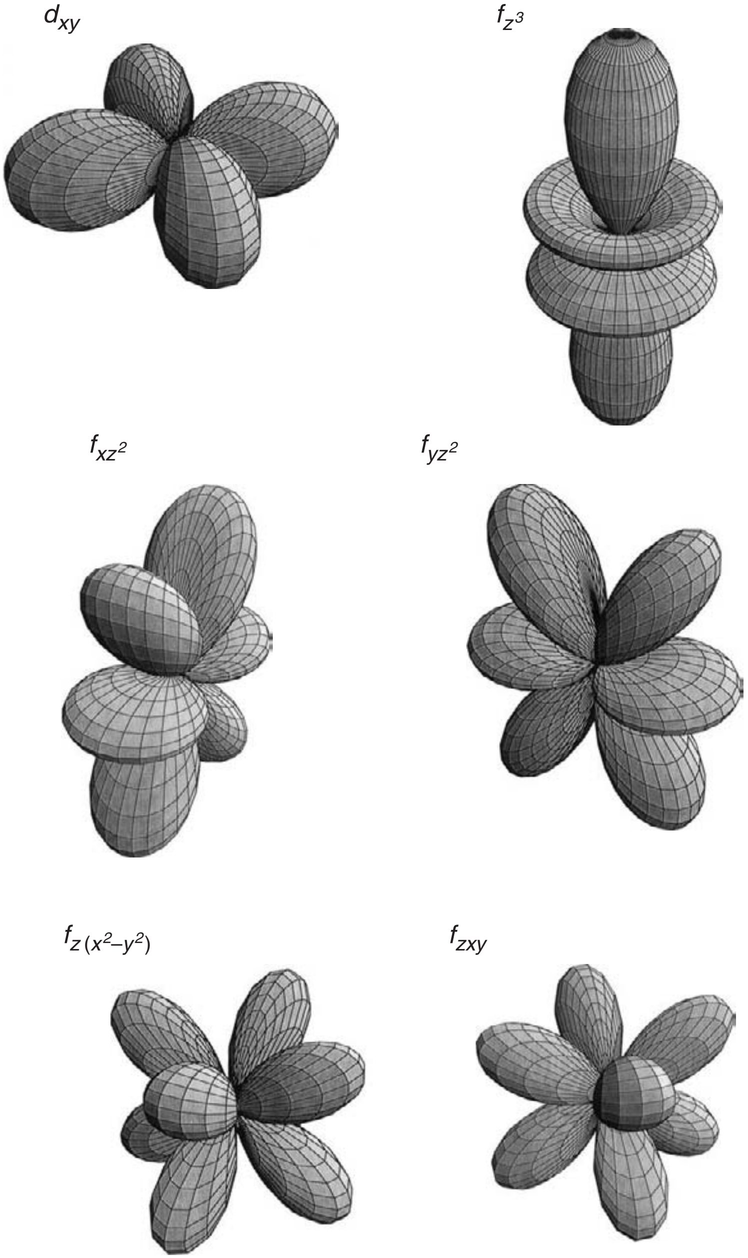


Fig. 1 (*continued*).



d_{xy}

f_{z^3}

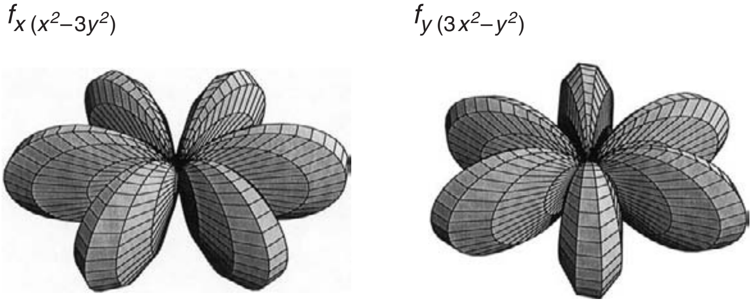
f_{xz^2}

f_{yz^2}

$f_{z(x^2-y^2)}$

f_{zxy}

Fig. 1 (*continued*).



$f_{x(x^2-3y^2)}$ $f_{y(3x^2-y^2)}$

Fig. 1 (*continued*).

Appendix IV: Radial Wavefunctions for Hydrogenlike Species

The normalized radial wavefunctions for hydrogenlike atoms can be expressed by

$$R_{n,\ell}(\rho) = -\sqrt{\frac{4(n-\ell-1)!}{n^4[(n+\ell)!]^3}}\left(\frac{Z}{a_0}\right)^{3/2}\left(\frac{2\rho}{n}\right)^{\ell} e^{-\rho/n}\mathscr{L}_{n+1}^{2\ell+1}(2\rho/n), \qquad (1)$$

with $\rho = Zr/a_0$, where r is the radial distance, Z is the atomic number and a_0 is the radius of the first Bohr orbit. The associated Laguerre polynomials, $\mathscr{L}_{n+1}^{2\ell+1}(2\rho/n)$, were introduced in Section 3.4.3. Here they have been multiplied by $(2\rho/n)^{\ell}e^{-\rho/n}$, which is the appropriate integrating factor. The radial functions $R_{n,\ell}(\rho)$ are given in Table 1 for the first three "shells" of hydrogenlike species.

Table 1 Some of the normalized radial wavefunctions for hydrogenlike species.

n	ℓ	symbol	normalizing factor	radial function
1	0	1s	$2(Z/a_0)^{3/2}$	$e^{-\rho}$
2	0	2s	$\frac{1}{2\sqrt{2}}(Z/a_0)^{3/2}$	$(2-\rho)e^{-\rho/2}$
	1	2p	$\frac{1}{2\sqrt{6}}(Z/a_0)^{3/2}$	$\rho e^{-\rho/2}$
3	0	3s	$\frac{2}{81\sqrt{3}}(Z/a_0)^{3/2}$	$(27-18\rho+2\rho^2)e^{-\rho/3}$
	1	3p	$\frac{4}{81\sqrt{6}}(Z/a_0)^{3/2}$	$\rho(6-\rho)e^{-\rho/3}$
	2	3d	$\frac{4}{81\sqrt{30}}(Z/a_0)^{3/2}$	$\rho^2 e^{-\rho/3}$

Appendix V: The Laplacian Operator in Spherical Coordinates

Spherical coordinates were introduced in Section 6.4. They were defined in Fig. 6-5 and by Eq. (6-54), namely,

$$x = r \sin\theta \cos\varphi, \tag{1}$$

$$y = r \sin\theta \sin\varphi \tag{2}$$

and

$$z = r \cos\theta. \tag{3}$$

Although transformations to various curvilinear coordinates can be carried out relatively easily with the use of the vector relations introduced in Section 5.15, it is often of interest to make the substitutions directly. Furthermore, it is a very good exercise in the manipulation of partial derivatives.

The relations given in Eq. (1) lead directly to the inverse expressions

$$r^2 = x^2 + y^2 + z^2, \tag{4}$$

$$\sin\theta = \frac{\sqrt{x^2 + y^2}}{r} \tag{5}$$

and

$$\tan\varphi = \frac{y}{x}. \tag{6}$$

The necessary derivatives can be evaluated from the above relations. For example, from Eq. (4)

$$\frac{\partial r}{\partial x} = \tfrac{1}{2}(x^2 + y^2 + z^2)^{-1/2}(2x) = \frac{x}{r} = \sin\theta \cos\varphi. \tag{7}$$

Similarly, Eq. (5) leads to

$$\frac{\partial \theta}{\partial x} = \frac{x \cos^2\theta}{z^2 \tan\theta} = \frac{\cos\theta \cos\varphi}{r} \tag{8}$$

and, Eq. (6) to

$$\frac{\partial\varphi}{\partial x} = -\frac{\sin\varphi}{r\sin\theta}. \tag{9}$$

Note that in the derivation of Eqs. (8) and (9) the relation $(d/d\varsigma)\tan\varsigma = \sec^2\varsigma$ has been employed, where ς can be identified with either θ or φ. The analogous derivatives can be easily derived by the same method. They are:

$$\frac{\partial r}{\partial y} = \sin\theta\sin\varphi, \tag{10}$$

$$\frac{\partial\theta}{\partial y} = \frac{\cos\theta\sin\varphi}{r}, \tag{11}$$

$$\frac{\partial\varphi}{\partial y} = \frac{\cos\varphi}{r\sin\theta}, \tag{12}$$

$$\frac{\partial r}{\partial z} = \cos\theta, \tag{13}$$

$$\frac{\partial\theta}{\partial z} = -\frac{\sin\theta}{r} \tag{14}$$

and

$$\frac{\partial\varphi}{\partial z} = 0. \tag{15}$$

The expressions for the various vector operators in spherical coordinates can be derived with the use of the chain rule. Thus, for example,

$$\begin{aligned}\frac{\partial}{\partial x} &= \left(\frac{\partial r}{\partial x}\right)\frac{\partial}{\partial r} + \left(\frac{\partial\theta}{\partial x}\right)\frac{\partial}{\partial\theta} + \left(\frac{\partial\varphi}{\partial x}\right)\frac{\partial}{\partial\varphi}\\ &= \sin\theta\cos\varphi\frac{\partial}{\partial r} + \frac{\cos\theta\cos\varphi}{r}\frac{\partial}{\partial\theta} - \frac{\sin\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi},\end{aligned} \tag{16}$$

with analogous relations for the two other operators. With the aid of these expressions the nabla, ∇, in spherical coordinates can be derived from Eq. (5-46).

To obtain the Laplacian in spherical coordinates it is necessary to take the appropriate second derivatives. Again, as an example, the derivative of Eq. (16) can be written as

$$
\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\right) = \frac{\partial^2}{\partial x^2} = \sin\theta\cos\varphi\Bigg[&\sin\theta\cos\varphi\frac{\partial^2}{\partial r^2} + \frac{\cos\theta\cos\varphi}{r}\frac{\partial^2}{\partial r\partial\theta} \\
&-\frac{\cos\theta\cos\varphi}{r^2}\frac{\partial}{\partial\theta} - \frac{\sin\varphi}{r\sin\theta}\frac{\partial^2}{\partial r\partial\varphi} + \frac{\sin\varphi}{r^2\sin\theta}\frac{\partial}{\partial\varphi}\Bigg] \\
&+\frac{\cos\theta\cos\varphi}{r}\Bigg[\sin\theta\cos\varphi\frac{\partial^2}{\partial r\partial\theta} + \cos\theta\cos\varphi\frac{\partial}{\partial r} \\
&+\frac{\cos\theta\cos\varphi}{r}\frac{\partial^2}{\partial\theta^2} - \frac{\sin\theta\cos\varphi}{r}\frac{\partial}{\partial\theta} - \frac{\sin\varphi}{r\sin\theta}\frac{\partial^2}{\partial\varphi\partial\theta} \\
&+\frac{\cos\theta\sin\varphi}{r\sin^2\theta}\frac{\partial}{\partial\varphi}\Bigg] - \frac{\sin\varphi}{r\sin\theta}\Bigg[\sin\theta\cos\varphi\frac{\partial^2}{\partial r\partial\varphi} \\
&-\sin\theta\sin\varphi\frac{\partial}{\partial r} + \frac{\cos\theta\cos\varphi}{r}\frac{\partial^2}{\partial\varphi\partial\theta} - \frac{\cos\theta\sin\varphi}{r}\frac{\partial}{\partial\theta} \\
&-\frac{\sin\varphi}{r\sin\theta}\frac{\partial^2}{\partial\varphi^2} - \frac{\cos\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi}\Bigg].
\end{aligned}
\tag{17}
$$

The corresponding operators in y and z are derived in the same way. The sum of these three operators yields the Laplacian as

$$
\begin{aligned}
\nabla^2 &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \\
&= \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2} + \frac{\cot\theta}{r^2}\frac{\partial}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}
\end{aligned}
\tag{18}
$$

or

$$
\nabla^2 = \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)\right] + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}. \tag{19}
$$

Equation (19) is the classic form of this operator in spherical coordinates as given in Eq. (6-55).

Appendix VI: The Divergence Theorem

The divergence theorem, usually attributed to Gauss, provides a relation between a volume V in space and the area S of the surface that bounds it. The theorem can be simply derived from the following argument.

Consider an element of V along the x direction that is bounded by the xy and xz planes, as shown in Fig. 1. The unit vectors $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are the outer normals with respect to the ends of the volume element shown. Thus, for any position vector $\boldsymbol{A}$ its components along a particular outer normal are given by $\boldsymbol{A} \cdot \boldsymbol{n}$. Furthermore, its components A_x along the x axis are functions of x. Thus,

$$\int_{x_1}^{x_2} \frac{\partial A_x}{\partial x}\,\mathrm{d}x = A_x\Big|_{x_1}^{x_2}, \tag{1}$$

where x_1 and x_2 are the values of x at which the element intersects the surface S. If the areas of the ends of the elements are $\mathrm{d}a_1$ and $\mathrm{d}a_2$, as indicated,

$$\mathrm{d}y\,\mathrm{d}z = -\mathrm{d}a_1\, cos(\boldsymbol{n}_1, x) + \mathrm{d}a_2\, cos(\boldsymbol{n}_2, x), \tag{2}$$

where $(\boldsymbol{n}_1, x)$ and $(\boldsymbol{n}_2, x)$ are the angles between the corresponding outer normals and the x axis.

If Eq. (1) is multiplied by $\mathrm{d}y\mathrm{d}z$, it becomes

$$\begin{aligned}\int_{x_1}^{x_2} \frac{\partial A_x}{\partial x}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z &= \left[A_{x_1}\, cos(\boldsymbol{n}_1, x)\,\mathrm{d}a_1 + A_{x_2}\, cos(\boldsymbol{n}_2, x)\,\mathrm{d}a_2\right] \\ &= A_x\, cos(\boldsymbol{n}, x)\,\mathrm{d}a, \end{aligned} \tag{3}$$

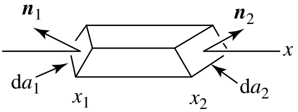


Fig. 1 A volume element in the x direction.

where $\mathrm{d}a = \mathrm{d}a_1 + \mathrm{d}a_2$. The summation of all elements in the x direction leads to

$$\iiint_v \frac{\partial A_x}{\partial x}\,\mathrm{d}v = \iint_S A_x \cos(\boldsymbol{n}, x)\,\mathrm{d}a, \tag{4}$$

with $\mathrm{d}v = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$. If this entire procedure is now repeated in the y and z directions, Eq. (4) can be generalized in the form

$$\iiint_v \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\right)\mathrm{d}v$$

$$= \iint_S \left[A_x \cos(\boldsymbol{n}, x) + A_y \cos(\boldsymbol{n}, y) + A_z \cos(\boldsymbol{n}, z)\right]\mathrm{d}a, \tag{5}$$

which can be written as

$$\iiint_v \nabla \cdot \boldsymbol{A}\,\mathrm{d}v = \iint_S \boldsymbol{A} \cdot \boldsymbol{n}\,\mathrm{d}a. \tag{6}$$

Equation (6) expresses the divergence theorem.

The divergence theorem has many applications. A very important case is that specified by Eq. (5-66), one of the four equations of Maxwell. It is specifically

$$\nabla \times \mathscr{H} = \boldsymbol{J} + \dot{\mathscr{D}}. \tag{7}$$

It leads directly to the equation of continuity for the charge density in a closed volume, *viz.*

$$\nabla \cdot \boldsymbol{J} = -\frac{\partial \rho}{\partial t}, \tag{8}$$

which is Eq. (5-68).

The equation of continuity also finds application in thermodynamics, as the flux density of heat from an enclosed volume must be compensated by a corresponding rate of temperature decrease within. Similarly, in fluid dynamics, if the volume contains an incompressible liquid, the flux density of flow from the volume results in an equivalent rate of decrease in the density within the enclosure.

Appendix VII: Determination of the Molecular Symmetry Group

A systematic approach to the determination of the point group that describes the symmetry of a molecular is suggested by the flow diagram of Fig. 1. If any difficulty is encountered in finding the symmetry group of a given structure, it is recommended that a molecular model be constructed – the familiar "sticks and stones" version that is employed in elementary courses in organic chemistry. It is then sufficient to follow the diagram and respond to the questions, as indicated. In general the answer is yes (**Y**) or no (**N**).

Play the symmetry game! Follow the flow diagram (Fig. 1). When you arrive at a point where there is a number in parenthesis, refer to the following comments:

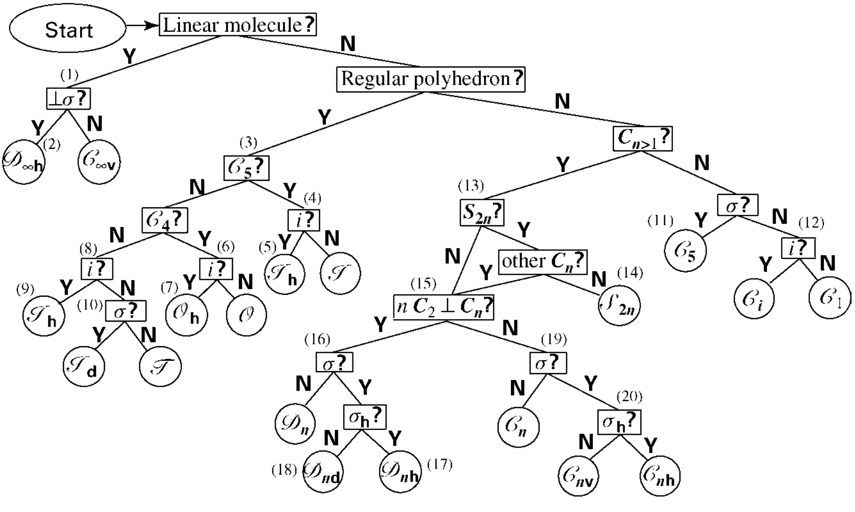


Fig. 1 Flow diagram for the determination of the symmetry group of a molecule.

(1) If the equilibrium structure of your molecule is linear, verify that it has a proper rotation axis of infinite order and an infinite number of planes of symmetry.

(2) If your molecule belongs to the group $\mathscr{D}_{\infty\mathbf{h}}$, it also has an infinite number of binary axes of rotation and a center of inversion. Please check.

(3) You say that your nonlinear molecule has the high symmetry of a regular polyhedron, such as a tetrahedron, cube, octahedron, dodecahedron, icosahedron, ... sphere. If it is a sphere, it is monatomic. On the other hand, if it is not monatomic, it has the symmetry of one of the Platonic solids (see the introduction to Chapter 8).

(4) You have replied that your molecule has a 5-fold axis of rotation. Verify that it also has 15 binary axes and ten ternary axes. Note that it belongs to one of the icosahedral groups. If you play soccer, consider the ball. Before you kick it, look at it. What is its symmetry?

(5) If your molecule has a center of inversion, as you have now indicated, it also must have 15 planes of symmetry. Can you find them?

(6) OK, your molecule does not have a C_5 axis. However, if it has a C_4 axis, it also has three binary rotation axes collinear with the C_4 and six other binary axes. Look carefully to be sure that your molecule indeed belongs to one of the octahedral groups.

(7) If your octahedral molecule has a center of symmetry, it also has nine planes of symmetry (three "horizontal" and six "diagonal"), as well as a number of improper rotation axes or orders four and six. Can you find all of them? If so, you can conclude that your molecule is of symmetry $\mathscr{O}_{\mathbf{h}}$.

(8) If there is neither a C_5 axis nor a C_4 axis, the symmetry of your molecule is that of one of the tetrahedral groups. Check that it also has four 3-fold and three binary rotation axes.

(9) If your molecule has a center of symmetry, it is of point group $\mathscr{T}_{\mathbf{h}}$.

(10) If, as in comment (8), your molecule has any planes of symmetry, it has six of them, as well as three improper axes of order four. If you find them, you can conclude that your molecule is of symmetry $\mathscr{T}_{\mathbf{d}}$. If not, it is of symmetry $\mathscr{T}$.

(11) You have replied that your molecule, that is not a regular polyhedron, does not have a proper rotation axis of order greater than one. If its only symmetry element is a plane, it belongs to the group $\mathscr{C}_{\mathbf{1h}} \equiv \mathscr{C}_{\mathbf{s}}$.

(12) However, if it has center of inversion, it belongs to the group $\mathscr{C}_{\mathbf{i}} \equiv \mathscr{S}_2$. It is then a very rare specimen. It is suggested that you repeat the analysis of its symmetry. On the other hand if your molecule does not have a center of inversion, its symmetry (or lack thereof) is described

by the group $\mathcal{C}_1$ and group theory cannot help you! Unfortunately, many molecules are in this category.

(13) You have specified the order of the proper rotation axis C_n is $n \geq 2$. Is there, then, an improper axis S_{2n}? (Note that if $n > 2$, the n-fold rotation axis C_n is by convention taken to be the vertical (z) axis).

(14) You have replied that there is indeed an axis S_{2n}. However, are there other binary axes C_2 perpendicular to the S_{2n}? If not, the symmetry of your molecule is described by one of the groups $\mathcal{S}_{2n}$ (Note that if n is odd, there is a center of inversion). However, this result is subject to doubt, as there are very few molecules of symmetry $\mathcal{S}_{2n}$.

(15) If you have arrived at this point, you have found no improper axis S_{2n}, but more than one proper axis of order n Or, you have identified at least one other axis of order n in addition to that collinear with S_{2n}. Now, look for n binary axes perpendicular to C_n. If you do not find any, the point group is one of the $\mathcal{C}$-type. Otherwise, there are n binary axes and the group is one of type $\mathcal{D}$.

(16) If there is no plane of symmetry, your molecule belongs to one of the $\mathcal{D}_{\mathbf{n}}$ groups.

(17) If there is a plane of symmetry perpendicular to the C_n axis, it is denoted by σ_h. Then, if your molecule is of symmetry $\mathcal{D}_{\mathbf{nh}}$, it also has n planes of symmetry in addition to the horizontal one. Furthermore, it must have an n-fold improper rotation axis (note that $i \equiv S_2$). In general if n is even, there is also a center of symmetry.

(18) If you have arrived at point group $\mathcal{D}_{\mathbf{nd}}$, your molecule must have n (diagonal) planes of symmetry in addition to the horizontal one. If n is odd, there is also a center of symmetry. The simplest case is that with $n = 2$. Look at the seams on a baseball or a tennis ball and verify that its symmetry is that of $\mathcal{D}_{2\mathbf{d}}$.

(19) The absence of n binary axes will lead you to one of the three $\mathcal{C}$-type point groups shown. Your molecule has the symmetry of one of the $\mathcal{C}_{\mathbf{n}}$ groups if it has no planes of symmetry.

(20) The presence or absence of a horizontal plane of symmetry will characterize the groups $\mathcal{C}_{\mathbf{nh}}$ or $\mathcal{C}_{\mathbf{nv}}$, respectively. To verify these possible results note that the former groups must have an improper rotation axis of order n ($\mathcal{C}_{\mathbf{1h}} \equiv \mathcal{C}_{\mathbf{s}}$). However, for the latter (groups $\mathcal{C}_{\mathbf{nv}}$), you will hopefully find n vertical planes, but no center of symmetry.

Appendix VIII: Character Tables for Some of the More Common Point Groups

Character tables are given in this appendix for the following point groups: (1) the groups which correspond to the 32 crystal classes (see Table 8-14), (2) groups containing 5-fold axes that may be needed to describe the symmetries of certain molecules or complex ions, and (3) the infinite groups $\mathscr{C}_{\infty\mathrm{v}}$ and $\mathscr{D}_{\infty\mathrm{h}}$ that are appropriate to linear structures.

Some comment on the notation for the irreducible representations (symmetry species) is necessary, as certain variations will be found in the literature. Molecular spectroscopists usually designate one-dimensional symmetry species by A or B, two-dimensional species by E, and three-dimensional species by F or T. Some authors prefer lower-case letters to capitals. The symmetry or antisymmetry of a given species with respect to the generating rotation operation distinguishes A and B. Subscripts 1 and 2 are used on A and B according to the symmetry with respect to rotation about a C_2 axis perpendicular to the generating axis. Symmetry with respect to inversion is indicated by a subscript g or u (see Section 4.4.2), while a prime or double prime identifies species that are, respectively, symmetric or antisymmetric with respect to a horizontal plane. Finally numerical subscripts are used to distinguish various doubly and triply degenerate species.

For linear molecules or ions the symbols are usually those derived from the term symbols for the electronic states of diatomic and other linear molecules. A capital Greek letter $\Sigma, \Pi, \Delta, \Phi, \ldots$ is used, corresponding to $\lambda = 0, 1, 2, 3, \ldots$, where λ is the quantum number for rotation about the molecular axis. For Σ species a superscript $+$ or $-$ is added to indicate the symmetry with respect to a plane that contains the molecular axis.

The components of the translation and rotation vectors are given as T_x, T_y, T_z and R_x, R_y, R_z, respectively. The components of the polarizability tensor appear as linear combinations such as $\alpha_{xx} + \alpha_{yy}$, *etc.*, that have the symmetry of the indicated irreducible representation.

Table 1 Character tables for the cyclic groups, $\mathcal{C}_n (n = 2, 3, 4, 5, 6)$.

$\mathcal{C}_2$	E	C_2		
A	1	1	T_z, R_z	$\alpha_{xx}, \alpha_{yy}, \alpha_{zz}, \alpha_{xy}$
B	1	-1	T_x, T_y, R_x, R_y	α_{yz}, α_{zx}

$\mathcal{C}_3$	E	C_3	C_3^2		$\varepsilon = exp(2\pi i/3)$
A	1	1	1	T_z, R_z	$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
E	1	ε	$\varepsilon^\bigstar$	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy}), (\alpha_{yz}, \alpha_{zx})$
	1	$\varepsilon^\bigstar$	ε		

$\mathcal{C}_4$	E	C_4	C_2	C_4^3		
A	1	1	1	1	T_z, R_z	$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
B	1	-1	1	-1		$\alpha_{xx} - \alpha_{yy}, \alpha_{zz}$
E	1	i	-1	$-i$	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz}, \alpha_{zx})$
	1	$-i$	-1	i		

$\mathcal{C}_5$	E	C_5	C_5^2	C_5^3	C_5^4		$\varepsilon = exp(2\pi i/5)$
A	1	1	1	1	1	T_z, R_z	$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
E_1	1	ε	ε^2	$\varepsilon^{2\bigstar}$	$\varepsilon^\bigstar$	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz}, \alpha_{zx})$
	1	$\varepsilon^\bigstar$	$\varepsilon^{2\bigstar}$	ε^2	ε		
E_2	1	ε^2	$\varepsilon^\bigstar$	ε	$\varepsilon^{2\bigstar}$		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
	1	$\varepsilon^{2\bigstar}$	ε	$\varepsilon^\bigstar$	ε^2		

$\mathcal{C}_6$	E	C_6	C_3	C_2	C_3^2	C_6^5		$\varepsilon = exp(2\pi i/6)$
A	1	1	1	1	1	1	T_z, R_z	$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
B	1	-1	1	-1	1	-1		
E_1	1	ε	$-\varepsilon^\bigstar$	-1	$-\varepsilon$	$\varepsilon^\bigstar$	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz} - \alpha_{zx})$
	1	$\varepsilon^\bigstar$	$-\varepsilon$	-1	$-\varepsilon^\bigstar$	ε		
E_2	1	$-\varepsilon^\bigstar$	$-\varepsilon$	1	$-\varepsilon^\bigstar$	$-\varepsilon$		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
	1	$-\varepsilon$	$-\varepsilon^\bigstar$	1	$-\varepsilon$	$-\varepsilon^\bigstar$		

Table 2 Character tables for the dihedral groups, $\mathscr{D}_n (n = 2, 3, 4, 5, 6)$.

$\mathscr{D}_2 = \mathscr{V}$	E	$C_2(z)$	$C_2(y)$	$C_2(x)$		
A	1	1	1	1		$\alpha_{xx}, \alpha_{yy}, \alpha_{zz}$
B$_1$	1	1	−1	−1	T_z, R_z	α_{xy}
B$_2$	1	−1	1	−1	T_y, R_y	α_{zx}
B$_3$	1	−1	−1	1	T_x, R_x	α_{yz}

$\mathscr{D}_3$	E	$2C_3$	$3C_2'$		
A$_1$	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A$_2$	1	1	−1	T_z, R_z	
E	2	−1	0	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy}), (\alpha_{yz}, \alpha_{zx})$

$\mathscr{D}_4$	E	$2C_4$	C_2	$2C_2'$	$2C_2''$		
A$_1$	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A$_2$	1	1	1	−1	−1	T_z, R_z	
B$_1$	1	−1	1	1	−1		$\alpha_{xx} - \alpha_{yy}$
B$_2$	1	−1	1	−1	1		α_{xy}
E	2	0	−2	0	0	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz}, \alpha_{zx})$

$\mathscr{D}_5$	E	$2C_5$	$2C_5^2$	$5C_2'$		$\alpha = 72°$
A$_1$	1	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A$_2$	1	1	1	−1	T_z, R_z	
E$_1$	2	$2\cos\alpha$	$2\cos 2\alpha$	0	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz}, \alpha_{zx})$
E$_2$	2	$2\cos 2\alpha$	$2\cos\alpha$	0		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$

$\mathscr{D}_6$	E	$2C_6$	$2C_3$	C_2	$3C_2'$	$3C_2''$		
A$_1$	1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A$_2$	1	1	1	1	−1	−1	T_z, R_z	
B$_1$	1	−1	1	−1	1	−1		
B$_2$	1	−1	1	−1	−1	1		
E$_1$	2	1	−1	−2	0	0	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz}, \alpha_{zx})$
E$_2$	2	−1	−1	2	0	0		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$

Table 3 Character tables for the groups $\mathscr{D}_{nh} (n = 2, 3, 4, 5, 6)$.

$\mathscr{D}_{2h} = \mathscr{V}_h$	E	$C_2(z)$	$C_2(y)$	$C_2(x)$	i	$\sigma(xy)$	$\sigma(xz)$	$\sigma(yz)$		
A_g	1	1	1	1	1	1	1	1		$\alpha_{xx}, \alpha_{yy}, \alpha_{zz}$
B_{1g}	1	1	−1	−1	1	1	−1	−1	R_z	α_{xy}
B_{2g}	1	−1	1	−1	1	−1	1	−1	R_y	α_{zx}
B_{3g}	1	−1	−1	1	1	−1	−1	1	R_x	α_{yz}
A_u	1	1	1	1	−1	−1	−1	−1		
B_{1u}	1	1	−1	−1	−1	−1	1	1	T_z	
B_{2u}	1	−1	1	−1	−1	1	−1	1	T_y	
B_{3u}	1	−1	−1	1	−1	1	1	−1	T_x	

$\mathscr{D}_{3h}$	E	$2C_3$	$3C_2'$	σ_h	$2S_3$	$3\sigma_v$		
A_1'	1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A_2'	1	1	−1	1	1	−1	R_z	
E'	2	−1	0	2	−1	0	(T_x, T_y)	$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
A_1''	1	1	1	−1	−1	−1		
A_2''	1	1	−1	−1	−1	1	T_z	
E''	2	−1	0	−2	1	0	(R_x, R_y)	$(\alpha_{yx}, \alpha_{zx})$

$\mathscr{D}_{4h}$	E	$2C_4$	C_2	$2C_2'$	$2C_2''$	i	$2S_4$	σ_h	$2\sigma_v$	$2\sigma_d$		
A_{1g}	1	1	1	1	1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A_{2g}	1	1	1	−1	−1	1	1	1	−1	−1	R_z	
B_{1g}	1	−1	1	1	−1	1	−1	1	1	−1		$\alpha_{xx} - \alpha_{yy}$
B_{2g}	1	−1	1	−1	1	1	−1	1	−1	1		α_{xy}
E_g	2	0	−2	0	0	2	0	−2	0	0	(R_x, R_y)	$(\alpha_{yz}, \alpha_{zx})$
A_{1u}	1	1	1	1	1	−1	−1	−1	−1	−1		
A_{2u}	1	1	1	−1	−1	−1	−1	−1	1	1	T_z	
B_{1u}	1	−1	1	1	−1	−1	1	−1	−1	1		
B_{2u}	1	−1	1	−1	1	−1	1	−1	1	−1		
E_u	2	0	−2	0	0	−2	0	2	0	0	(T_x, T_y)	

$\mathscr{D}_{5h}$	E	$2C_5$	$2C_5^2$	$5C_2'$	σ_h	$2S_5$	$2S_5^3$	$5\sigma_v$		$\alpha = 72°$
A_1'	1	1	1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A_2'	1	1	1	−1	1	1	1	−1	R_z	
E_1'	2	$2\cos\alpha$	$2\cos 2\alpha$	0	2	$2\cos\alpha$	$2\cos 2\alpha$	0	(T_x, T_y)	
E_2'	2	$2\cos 2\alpha$	$2\cos\alpha$	0	2	$2\cos 2\alpha$	$2\cos\alpha$	0		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
A_1''	1	1	1	1	−1	−1	−1	−1		
A_2''	1	1	1	−1	−1	−1	−1	1	T_z	
E_1''	2	$2\cos\alpha$	$2\cos 2\alpha$	0	−2	$-2\cos\alpha$	$-2\cos 2\alpha$	0	(R_x, R_y)	$(\alpha_{yz}, \alpha_{zx})$
E_2''	2	$2\cos 2\alpha$	$2\cos\alpha$	0	−2	$-2\cos 2\alpha$	$-2\cos\alpha$	0		

Table 3 (*Continued*).

$\mathcal{D}_{6h}$	E	$2C_6$	$2C_3$	C_2	$3C_2'$	$3C_2''$	i	$2S_3$	$2S_6$	σ_h	$3\sigma_d$	$3\sigma_v$		
A_{1g}	1	1	1	1	1	1	1	1	1	1	1	1		$\alpha_{xx}+\alpha_{yy}, \alpha_{zz}$
A_{2g}	1	1	1	1	−1	−1	1	1	1	1	−1	−1	R_z	
B_{1g}	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1		
B_{2g}	1	−1	1	−1	−1	1	1	−1	1	−1	−1	1		
E_{1g}	2	1	−1	−2	0	0	2	1	−1	−2	0	0	(R_x, R_y)	$(\alpha_{yz}, \alpha_{zx})$
E_{2g}	2	−1	−1	2	0	0	2	−1	−1	2	0	0		$(\alpha_{xx}-\alpha_{yy}, \alpha_{xy})$
A_{1u}	1	1	1	1	1	1	−1	−1	−1	−1	−1	−1		
A_{2u}	1	1	1	1	−1	−1	−1	−1	−1	−1	1	1	T_z	
B_{1u}	1	−1	1	−1	1	−1	−1	1	−1	1	−1	1		
B_{2u}	1	−1	1	−1	−1	1	−1	1	−1	1	1	−1		
E_{1u}	2	1	−1	−2	0	0	−2	−1	1	2	0	0	(T_x, T_y)	
E_{2u}	2	−1	−1	2	0	0	−2	1	1	−2	0	0		

Table 4 Character tables for the groups $\mathcal{S}_n (n = 2, 4, 6, 8)$.

$\mathcal{S}_2 \equiv \mathcal{C}_i$	E	i		
A_g	1	1	R_x, R_y, R_z	$\alpha_{xx}, \alpha_{yy}, \alpha_{zz}, \alpha_{xy}$ α_{yz}, α_{zx}
A_u	1	−1	T_x, T_y, T_z	

$\mathcal{S}_4$	E	S_4	C_2	S_4^3		
A	1	1	1	1	R_z	$\alpha_{xx}+\alpha_{yy}, \alpha_{zz}$
B	1	−1	1	−1	T_z	$\alpha_{xx}-\alpha_{yy}, \alpha_{xy}$
E	1	i	−1	$-i$	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz}, \alpha_{zx})$
	1	$-i$	−1	i		

$\mathcal{S}_6$	E	C_3	C_3^2	i	S_6^5	S_6		$\varepsilon = exp(2\pi i/3)$
A_g	1	1	1	1	1	1	R_z	$\alpha_{xx}+\alpha_{yy}, \alpha_{zz}$
E_g	1	ε	$\varepsilon^\star$	1	ε	$\varepsilon^\star$	(R_x, R_y)	$(\alpha_{xx}-\alpha_{yy}, \alpha_{xy}), (\alpha_{yz}, \alpha_{zx})$
	1	$\varepsilon^\star$	ε	1	$\varepsilon^\star$	ε		
A_u	1	1	1	−1	−1	−1	T_z	
E_u	1	ε	$\varepsilon^\star$	−1	$-\varepsilon$	$-\varepsilon^\star$	(T_x, T_y)	
	1	$\varepsilon^\star$	ε	−1	$-\varepsilon^\star$	$-\varepsilon$		

$\mathcal{S}_8$	E	S_8	C_4	S_8^3	C_2	S_8^5	C_4^3	S_8^7	$\varepsilon = exp(2\pi i/8)$	
A	1	1	1	1	1	1	1	1	R_z	$\alpha_{xx}+\alpha_{yy}, \alpha_{zz}$
B	1	−1	1	−1	1	−1	1	−1	T_z	
E_1	1	ε	i	$-\varepsilon^\star$	−1	$-\varepsilon$	$-i$	$\varepsilon^\star$	(T_x, T_y)	
	1	$\varepsilon^\star$	$-i$	$-\varepsilon$	−1	$-\varepsilon^\star$	i	ε		
E_2	1	i	−1	$-i$	1	i	−1	$-i$		$(\alpha_{xx}-\alpha_{yy}, \alpha_{xy})$
	1	$-i$	−1	i	1	$-i$	−1	i		
E_3	1	$-\varepsilon^\star$	$-i$	ε	−1	$\varepsilon^\star$	i	$-\varepsilon$	(R_x, R_y)	$(\alpha_{yz}, \alpha_{zx})$
	1	$-\varepsilon$	i	$\varepsilon^\star$	−1	ε	$-i$	$-\varepsilon^\star$		

Table 5 Character tables for the groups $\mathscr{C}_{nh}(n = 1, 2, 3, 4, 5, 6)$.

$\mathscr{C}_{1h} \equiv \mathscr{C}_s$	E	σ_h		
A′	1	1	T_x, T_y, R_z	$\alpha_{xx}, \alpha_{yy}, \alpha_{zz}, \alpha_{xy}$
A″	1	−1	T_z, R_x, R_y	α_{yz}, α_{zx}

$\mathscr{C}_{2h}$	E	C_2	i	σ_h		
A_g	1	1	1	1	R_z	$\alpha_{xx}, \alpha_{yy}, \alpha_{zz}, \alpha_{xy}$
B_g	1	−1	1	−1	R_x, R_y	α_{yz}, α_{zx}
A_u	1	1	−1	−1	T_z	
B_u	1	−1	−1	1	T_x, T_y	

$\mathscr{C}_{3h}$	E	C_3	C_3^2	σ_h	S_3	S_3^5		$\varepsilon = exp(2\pi i/3)$
A′	1	1	1	1	1	1	R_z	$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
E′	1	ε	$\varepsilon^{\bigstar}$	1	ε	$\varepsilon^{\bigstar}$	(T_x, T_y)	$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
	1	$\varepsilon^{\bigstar}$	ε	1	$\varepsilon^{\bigstar}$	ε		
A″	1	1	1	−1	−1	−1	T_z	
E″	1	ε	$\varepsilon^{\bigstar}$	−1	$-\varepsilon$	$-\varepsilon^{\bigstar}$	(R_x, R_y)	$(\alpha_{yz}, \alpha_{zx})$
	1	$\varepsilon^{\bigstar}$	ε	−1	$-\varepsilon^{\bigstar}$	$-\varepsilon$		

$\mathscr{C}_{4h}$	E	C_4	C_2	C_4^3	i	S_4^3	σ_h	S_4		
A_g	1	1	1	1	1	1	1	1	R_z	$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
B_g	1	−1	1	−1	1	−1	1	−1		$\alpha_{xx} - \alpha_{yy}, \alpha_{xy}$
E_g	1	i	−1	$-i$	1	i	−1	$-i$	(R_x, R_y)	$(\alpha_{yz}, \alpha_{zx})$
	1	$-i$	−1	i	1	$-i$	−1	i		
A_u	1	1	1	1	−1	−1	−1	−1	T_z	
B_u	1	−1	1	−1	−1	1	−1	1		
E_u	1	i	−1	$-i$	−1	$-i$	1	i	(T_x, T_y)	
	1	$-i$	−1	i	−1	i	1	$-i$		

$\mathscr{C}_{5h}$	E	C_5	C_5^2	C_5^3	C_5^4	σ_h	S_5	S_5^7	S_5^3	S_5^9		$\varepsilon = exp(2\pi i/5)$
A′	1	1	1	1	1	1	1	1	1	1	R_z	$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
E_1'	1	ε	ε^2	$\varepsilon^{2\bigstar}$	$\varepsilon^{\bigstar}$	1	ε	ε^2	$\varepsilon^{2\bigstar}$	$\varepsilon^{\bigstar}$	(T_z, T_y)	
	1	$\varepsilon^{\bigstar}$	$\varepsilon^{2\bigstar}$	ε^2	ε	1	$\varepsilon^{\bigstar}$	$\varepsilon^{2\bigstar}$	ε^2	ε		
E_2'	1	ε^2	$\varepsilon^{\bigstar}$	ε	$\varepsilon^{2\bigstar}$	1	ε^2	$\varepsilon^{\bigstar}$	ε	$\varepsilon^{2\bigstar}$		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
	1	$\varepsilon^{2\bigstar}$	ε	$\varepsilon^{\bigstar}$	ε^2	1	$\varepsilon^{2\bigstar}$	ε	$\varepsilon^{\bigstar}$	ε^2		
A″	1	1	1	1	1	−1	−1	−1	−1	−1	T_z	
E_1''	1	ε	ε^2	$\varepsilon^{2\bigstar}$	$\varepsilon^{\bigstar}$	−1	$-\varepsilon$	$-\varepsilon^2$	$-\varepsilon^{2\bigstar}$	$-\varepsilon^{\bigstar}$	(R_x, R_y)	$(\alpha_{yz} + \alpha_{zx})$
	1	$\varepsilon^{\bigstar}$	$\varepsilon^{2\bigstar}$	ε^2	ε	−1	$-\varepsilon^{\bigstar}$	$-\varepsilon^{2\bigstar}$	$-\varepsilon^2$	$-\varepsilon$		
E_2''	1	ε^2	$\varepsilon^{\bigstar}$	ε	$\varepsilon^{2\bigstar}$	−1	$-\varepsilon^2$	$-\varepsilon^{\bigstar}$	$-\varepsilon$	$-\varepsilon^{2\bigstar}$		
	1	$\varepsilon^{2\bigstar}$	ε	$\varepsilon^{\bigstar}$	ε^2	−1	$-\varepsilon^{2\bigstar}$	$-\varepsilon$	$-\varepsilon^{\bigstar}$	$-\varepsilon^2$		

Table 5 (*Continued*).

$\mathscr{C}_{6h}$	E	C_6	C_3	C_2	C_3^2	C_6^5	i	S_3^5	S_6^5	σ_h	S_6	S_3		$\varepsilon = exp(2\pi i/6)$
A_g	1	1	1	1	1	1	1	1	1	1	1	1	R_z	$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
B_g	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1		
E_{1g}	1	ε	$-\varepsilon^\star$	−1	$-\varepsilon$	$\varepsilon^\star$	1	ε	$-\varepsilon^\star$	−1	$-\varepsilon$	$\varepsilon^\star$	(R_x, R_y)	$(\alpha_{yz}, \alpha_{zx})$
	1	$\varepsilon^\star$	$-\varepsilon$	−1	$-\varepsilon^\star$	ε	1	$\varepsilon^\star$	$-\varepsilon$	−1	$-\varepsilon^\star$	ε		
E_{2g}	1	$-\varepsilon^\star$	$-\varepsilon$	1	$-\varepsilon^\star$	$-\varepsilon$	1	$-\varepsilon^\star$	$-\varepsilon$	1	$-\varepsilon^\star$	$-\varepsilon$		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
	1	$-\varepsilon$	$-\varepsilon^\star$	1	$-\varepsilon$	$-\varepsilon^\star$	1	$-\varepsilon$	$-\varepsilon^\star$	1	$-\varepsilon$	$-\varepsilon^\star$		
A_u	1	1	1	1	1	1	−1	−1	−1	−1	−1	−1	T_z	
B_u	1	−1	1	−1	1	−1	−1	1	−1	1	−1	1		
E_{1u}	1	ε	$-\varepsilon^\star$	−1	$-\varepsilon$	$\varepsilon^\star$	−1	$-\varepsilon$	$\varepsilon^\star$	1	ε	$-\varepsilon^\star$	(T_x, T_y)	
	1	$\varepsilon^\star$	$-\varepsilon$	−1	$-\varepsilon^\star$	ε	−1	$-\varepsilon^\star$	ε	1	$\varepsilon^\star$	$-\varepsilon$		
E_{2u}	1	$-\varepsilon^\star$	$-\varepsilon$	1	$-\varepsilon^\star$	$-\varepsilon$	−1	$\varepsilon^\star$	ε	−1	$\varepsilon^\star$	ε		
	1	$-\varepsilon$	$-\varepsilon^\star$	1	$-\varepsilon$	$-\varepsilon^\star$	−1	ε	$\varepsilon^\star$	−1	ε	$\varepsilon^\star$		

Table 6 Character tables for the groups $\mathscr{C}_{nv} (n = 2, 3, 4, 5, 6)$.

$\mathscr{C}_{2v}$	E	C_2	$\sigma_v(xz)$	$\sigma_v(yz)$		
A_1	1	1	1	1	T_z	$\alpha_{xx}, \alpha_{yy}, \alpha_{zz}$
A_2	1	1	−1	−1	R_z	α_{xy}
B_1	1	−1	1	−1	T_x, R_y	α_{zx}
B_2	1	−1	−1	1	T_y, R_x	α_{yz}

$\mathscr{C}_{3v}$	E	$2C_3$	$3\sigma_v$		
A_1	1	1	1	T_z	$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A_2	1	1	−1	R_z	
E	2	−1	0	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy}), (\alpha_{yz}, \alpha_{zx})$

$\mathscr{C}_{4v}$	E	$2C_4$	C_2	$2\sigma_v$	$2\sigma_d$		
A_1	1	1	1	1	1	T_x	$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A_2	1	1	1	−1	−1	R_z	
B_1	1	−1	1	1	−1		$\alpha_{xx} - \alpha_{yy}$
B_2	1	−1	1	−1	1		α_{xy}
E	2	0	−2	0	0	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz}, \alpha_{zx})$

$\mathscr{C}_{5v}$	E	$2C_5$	$2C_5^2$	$5\sigma_v$		$\alpha = 72°$
A_1	1	1	1	1	T_x	$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A_2	1	1	1	−1	R_z	
E_1	2	$2\cos\alpha$	$2\cos 2\alpha$	0	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz}, \alpha_{zx})$
E_2	2	$2\cos 2\alpha$	$2\cos\alpha$	0		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$

$\mathscr{C}_{6v}$	E	$2C_6$	$2C_3$	C_2	$3\sigma_v$	$3\sigma_d$		
A_1	1	1	1	1	1	1	T_z	$\alpha_{xx} + \alpha_{yy}, \alpha_{xy}$
A_2	1	1	1	1	−1	−1	R_z	
B_1	1	−1	1	−1	1	−1		
B_2	1	−1	1	−1	−1	1		
E_1	2	1	−1	−2	0	0	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz}, \alpha_{zx})$
E_2	2	−1	−1	2	0	0		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$

Table 7 Character tables for the groups, $\mathscr{D}_{nd} (n = 2, 3, 4, 5, 6)$.

$\mathscr{D}_{2d} = \mathscr{V}_d$	E	$2S_4$	C_2	$2C_2'$	$2\sigma_d$		
A_1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A_2	1	1	1	−1	−1	R_z	
B_1	1	−1	1	1	−1		$\alpha_{xx} - \alpha_{yy}$
B_2	1	−1	1	−1	1	T_z	α_{xy}
E	2	0	−2	0	0	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz}, \alpha_{zx})$

$\mathscr{D}_{3d}$	E	$2C_3$	$3C_2$	i	$2S_6$	$3\sigma_d$		
A_{1g}	1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A_{2g}	1	1	−1	1	1	−1	R_z	
E_g	2	−1	0	2	−1	0	(R_x, R_y)	$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy}), (\alpha_{yz}, \alpha_{zx})$
A_{1u}	1	1	1	−1	−1	−1		
A_{2u}	1	1	−1	−1	−1	1	T_z	
E_u	2	−1	0	−2	1	0	(T_x, T_y)	

$\mathscr{D}_{4d}$	E	$2S_8$	$2C_4$	$2S_8^3$	C_2	$4C_2'$	$4\sigma_d$		
A_1	1	1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A_2	1	1	1	1	1	−1	−1	R_z	
B_1	1	−1	1	−1	1	1	−1		
B_2	1	−1	1	−1	1	−1	1	T_z	
E_1	2	$\sqrt{2}$	0	$-\sqrt{2}$	−2	0	0	(T_x, T_y)	
E_2	2	0	−2	0	2	0	0		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
E_3	2	$-\sqrt{2}$	0	$\sqrt{2}$	−2	0	0	(R_x, R_y)	$(\alpha_{yz}, \alpha_{zx})$

$\mathscr{D}_{5d}$	E	$2C_5$	$2C_5^2$	$5C_2$	i	$2S_{10}^3$	$2S_{10}$	$5\sigma_d$		$\alpha = 72°$
A_{1g}	1	1	1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A_{2g}	1	1	1	−1	1	1	1	−1	R_z	
E_{1g}	2	$2\cos\alpha$	$2\cos 2\alpha$	0	2	$2\cos\alpha$	$2\cos 2\alpha$	0	(R_x, R_y)	$(\alpha_{yz}, \alpha_{zx})$
E_{2g}	2	$2\cos 2\alpha$	$2\cos\alpha$	0	2	$2\cos 2\alpha$	$2\cos\alpha$	0		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
A_{1u}	1	1	1	1	−1	−1	−1	−1		
A_{2u}	1	1	1	−1	−1	−1	−1	1	T_z	
E_{1u}	2	$2\cos\alpha$	$2\cos 2\alpha$	0	−2	$-2\cos\alpha$	$-2\cos 2\alpha$	0	(T_x, T_y)	
E_{2u}	2	$2\cos 2\alpha$	$2\cos\alpha$	0	−2	$-2\cos 2\alpha$	$-2\cos\alpha$	0		

$\mathscr{D}_{6d}$	E	$2S_{12}$	$2C_6$	$2S_4$	$2C_3$	$2S_{12}^5$	C_2	$6C_2'$	$6\sigma_d$		
A_1	1	1	1	1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
A_2	1	1	1	1	1	1	1	−1	−1	R_z	
B_1	1	−1	1	−1	1	−1	1	1	−1		
B_2	1	−1	1	−1	1	−1	1	−1	1	T_z	
E_1	2	$\sqrt{3}$	1	0	−1	$-\sqrt{3}$	−2	0	0	(T_x, T_y)	
E_2	2	1	−1	−2	−1	1	2	0	0		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
E_3	2	0	−2	0	2	0	−2	0	0		
E_4	2	−1	−1	2	−1	−1	2	0	0		
E_5	2	$-\sqrt{3}$	1	0	−1	$\sqrt{3}$	−2	0	0	(R_x, R_y)	$(\alpha_{yz}, \alpha_{zx})$

Table 8 Character tables for the cubic groups.

$\mathscr{T}$	E	$4C_3$	$4C_3^2$	$3C_2$		$\varepsilon = exp(2\pi i/3)$
A	1	1	1	1		$\alpha_{xx} + \alpha_{yy} + \alpha_{zz}$
E	1	ε	$\varepsilon^\star$	1		$(\alpha_{xx} + \alpha_{yy} - 2\alpha_{zz}, \alpha_{xx} - \alpha_{yy})$
	1	$\varepsilon^\star$	ε	1		
T ≡ F	3	0	0	−1	**T**, **R**	$(\alpha_{xy}, \alpha_{yz}, \alpha_{zx})$

$\mathscr{T}_h$	E	$4C_3$	$4C_3^2$	$3C_2$	i	$4S_6$	$4S_6^5$	3σ		$\varepsilon = exp(2\pi i/3)$
A_g	1	1	1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy} + \alpha_{zz}$
E_g	1	ε	$\varepsilon^\star$	1	1	$\varepsilon^\star$	ε	1		$(\alpha_{xx} + \alpha_{yy} - 2\alpha_{zz},$
	1	$\varepsilon^\star$	ε	1	1	ε	$\varepsilon^\star$	1		$\alpha_{xx} - \alpha_{yy})$
$T_g \equiv F_g$	3	0	0	−1	3	0	0	−1	**R**	$(\alpha_{xy}, \alpha_{yz}, \alpha_{zx})$
A_u	1	1	1	1	−1	−1	−1	−1		
E_u	1	ε	$\varepsilon^\star$	1	−1	$-\varepsilon^\star$	$-\varepsilon$	−1		
	1	$\varepsilon^\star$	ε	1	−1	$-\varepsilon$	$-\varepsilon^\star$	−1		
$T_u \equiv F_u$	3	0	0	−1	−3	0	0	1	**T**	

$\mathscr{T}_d$	E	$8C_3$	$3C_2$	$6S_4$	$6\sigma_d$		
A_1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy} + \alpha_{zz}$
A_2	1	1	1	−1	−1		
E	2	−1	2	0	0		$(\alpha_{xx} + \alpha_{yy} - 2\alpha_{zz}, \alpha_{xx} - \alpha_{yy})$
$T_1 \equiv F_1$	3	0	−1	1	−1	**R**	
$T_2 \equiv F_2$	3	0	−1	−1	1	**T**	$(\alpha_{xy}, \alpha_{yz}, \alpha_{zx})$

$\mathscr{O}_h$	E	$8C_3$	$3C_2$	$6C_4$	$6C_2'$	i	$8S_6$	$3\sigma_h$	$6S_4$	$6\sigma_d$		
A_{1g}	1	1	1	1	1	1	1	1	1	1		$\alpha_{xx} + \alpha_{yy} + \alpha_{zz}$
A_{2g}	1	1	1	−1	−1	1	1	1	−1	−1		
E_g	2	−1	2	0	0	2	−1	2	0	0		$(\alpha_{xx} + \alpha_{yy} - 2\alpha_{zz},$ $\alpha_{xx} - \alpha_{yy})$
$T_{1g} \equiv F_{1g}$	3	0	−1	1	−1	3	0	−1	1	−1	**R**	
$T_{2g} \equiv F_{2g}$	3	0	−1	−1	1	3	0	−1	−1	1		$(\alpha_{xy}, \alpha_{yz}, \alpha_{zx})$
A_{1u}	1	1	1	1	1	−1	−1	−1	−1	−1		
A_{2u}	1	1	1	−1	−1	−1	−1	−1	1	1		
E_u	2	−1	2	0	0	−2	1	−2	0	0		
$T_{1u} \equiv F_{1u}$	3	0	−1	1	−1	−3	0	1	−1	1	**T**	
$T_{2u} \equiv F_{2u}$	3	0	−1	−1	1	−3	0	1	1	−1		

Table 9 Character tables for the icosahedral groups.

$\mathscr{I}$	E	$12C_5$	$12C_5^2$	$20C_3$	$15C_2$		
A	1	1	1	1	1		$\alpha_{zz}+\alpha_{yy}+\alpha_{zz}$
$T_1 \equiv F_1$	3	$\frac{1+\sqrt{5}}{2}$	$\frac{1-\sqrt{5}}{2}$	0	−1	**T**, **R**	
$T_2 \equiv F_2$	3	$\frac{1-\sqrt{5}}{2}$	$\frac{1+\sqrt{5}}{2}$	0	−1		
G	4	−1	−1	1	0		
H	5	0	0	−1	1		$(\alpha_{xx}+\alpha_{yy}-2\alpha_{zz}, \alpha_{xx}-\alpha_{yy}, \alpha_{xy}, \alpha_{yz}, \alpha_{zx})$

$\mathscr{I}_h$	E	$12C_5$	$12C_5^2$	$20C_3$	$15C_2$	i	$12S_{10}$	$12S_{10}^3$	$20S_6$	15σ		
A_g	1	1	1	1	1	1	1	1	1	1		$\alpha_{xx}+\alpha_{yy}+\alpha_{zz}$
$T_{1g} \equiv F_{1g}$	3	$\frac{1}{2}(1+\sqrt{5})$	$\frac{1}{2}(1-\sqrt{5})$	0	−1	3	$\frac{1}{2}(1-\sqrt{5})$	$\frac{1}{2}(1+\sqrt{5})$	0	−1	**R**	
$T_{2g} \equiv F_{2g}$	3	$\frac{1}{2}(1-\sqrt{5})$	$\frac{1}{2}(1+\sqrt{5})$	0	−1	3	$\frac{1}{2}(1+\sqrt{5})$	$\frac{1}{2}(1-\sqrt{5})$	0	−1		
G_g	4	−1	−1	1	0	4	−1	−1	1	0		
H_g	5	0	0	−1	1	5	0	0	−1	1		$(\alpha_{xx}+\alpha_{yy}-2\alpha_{zz},$ $\alpha_{xx}-\alpha_{yy}, \alpha_{xy},$ $\alpha_{yz}, \alpha_{zx})$
A_u	1	1	1	1	1	−1	−1	−1	−1	−1		
$T_{1u} \equiv F_{1u}$	3	$\frac{1}{2}(1+\sqrt{5})$	$\frac{1}{2}(1-\sqrt{5})$	0	−1	−3	$-\frac{1}{2}(1-\sqrt{5})$	$-\frac{1}{2}(1+\sqrt{5})$	0	1	**T**	
$T_{2u} \equiv F_{2u}$	3	$\frac{1}{2}(1-\sqrt{5})$	$\frac{1}{2}(1+\sqrt{5})$	0	−1	−3	$-\frac{1}{2}(1+\sqrt{5})$	$-\frac{1}{2}(1-\sqrt{5})$	0	1		
G_u	4	−1	−1	1	0	−4	1	1	−1	0		
H_u	5	0	0	−1	1	−5	0	0	1	−1		

Table 10 Character tables for the infinite groups of linear structures.

$\mathscr{C}_{\infty v}$	E	$2C_{\infty}^{\phi}$	$\cdots$	$\infty\sigma_v$		
$A_1 \equiv \Sigma^+$	1	1	$\cdots$	1	T_z	$\alpha_{zz} + \alpha_{yy}, \alpha_{zz}$
$A_2 \equiv \Sigma^-$	1	1	$\cdots$	-1	R_z	
$E_1 \equiv \Pi$	2	$2\cos\phi$	$\cdots$	0	$(T_x, T_y), (R_x, R_y)$	$(\alpha_{yz}, \alpha_{zx})$
$E_2 \equiv \Delta$	2	$2\cos 2\phi$	$\cdots$	0		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
$E_3 \equiv \Phi$	2	$2\cos 3\phi$	$\cdots$	0		
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$		

$\mathscr{D}_{\infty h}$	E	$2C_{\infty}^{\phi}$	$\cdots$	$\infty\sigma_v$	i	$2S_{\infty}^{\phi}$	$\cdots$	∞C_2		
$\Sigma_g{}^+$	1	1	$\cdots$	1	1	1	$\cdots$	1		$\alpha_{xx} + \alpha_{yy}, \alpha_{zz}$
$\Sigma_g{}^-$	1	1	$\cdots$	-1	1	1	$\cdots$	-1	R_z	
Π_g	2	$2\cos\phi$	$\cdots$	0	2	$-2\cos\phi$	$\cdots$	0	(R_z, R_y)	$(\alpha_{yz}, \alpha_{zx})$
Δ_g	2	$2\cos 2\phi$	$\cdots$	0	2	$2\cos 2\phi$	$\cdots$	0		$(\alpha_{xx} - \alpha_{yy}, \alpha_{xy})$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$		
Σ_u^+	1	1	$\cdots$	1	-1	-1	$\cdots$	-1	T_z	
Σ_u^-	1	1	$\cdots$	-1	-1	-1	$\cdots$	1		
Π_u	2	$2\cos\phi$	$\cdots$	0	-2	$2\cos\phi$	$\cdots$	0	(T_x, T_y)	
Δ_u	2	$2\cos 2\phi$	$\cdots$	0	-2	$-2\cos 2\phi$	$\cdots$	0		
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$		

Appendix IX: Matrix Elements for the Harmonic Oscillator

Some of the more useful matrix elements for the harmonic oscillator are presented in the following table. They are given as functions of the dimensionless quantities $\xi = 2\pi x\sqrt{\nu m/h}$ and $\sigma = 2\varepsilon/h\nu$, as defined in Section 6.2.

$$\langle v|\xi|v+1\rangle = \sqrt{\tfrac{1}{2}(v+1)}$$

$$\langle v|\xi|v-1\rangle = \sqrt{\tfrac{1}{2}v}$$

$$\langle v|\xi|v'\rangle = 0, \text{ if } v' \neq v \pm 1$$

$$\langle v|\xi^2|v+2\rangle = \tfrac{1}{2}\sqrt{(v+1)(v+2)}$$

$$\langle v|\xi^2|v\rangle = v + \tfrac{1}{2}$$

$$\langle v|\xi^2|v-2\rangle = \tfrac{1}{2}\sqrt{v(v-1)}$$

$$\langle v|\xi^2|v'\rangle = 0, \text{ if } v' \neq v \pm 2$$

$$\langle v|\xi^3|v+3\rangle = \tfrac{1}{2}\sqrt{(v+1)(v+2)(v+3)/2}$$

$$\langle v|\xi^3|v+1\rangle = \tfrac{3}{2}\sqrt{[(v+1)/2]^3}$$

$$\langle v|\xi^3|v-1\rangle = \tfrac{3}{2}\sqrt{v^3/2}$$

$$\langle v|\xi^3|v-3\rangle = \tfrac{1}{2}\sqrt{v(v-1)(v-2)/2}$$

$$\langle v|\xi^3|v'\rangle = 0, \text{ if } v' \neq v \pm 1 \text{ or } v' \neq v \pm 3$$

$$\langle v|\xi^4|v+4\rangle = \tfrac{1}{4}\sqrt{(v+1)(v+2)(v+3)(v+4)/2}$$

$$\langle v|\xi^4|v+2\rangle = \tfrac{1}{2}(2v+3)\sqrt{(v+1)(v+2)}$$

$$\langle v|\xi^4|v\rangle = \tfrac{3}{4}(2v^2+2v+1)$$

$$\langle v|\xi^4|v-2\rangle = \tfrac{1}{2}(2v-1)\sqrt{v(v-1)}$$

$$\langle v|\xi^4|v-4\rangle = \tfrac{1}{4}\sqrt{v(v-1)(v-2)(v-3)/2}$$

$$\langle v|\xi^4|v'\rangle = 0, \text{ if } v' \neq v,\ v \pm 2 \text{ or } v \pm 4$$

Appendix X: Further Reading

Applied mathematics

Abramowitz, Milton and Stegun, Irene A. (eds.) *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, National Bureau of Standards, Applied Mathematics Series 55, U.S. Government Printing Office, Washington, D.C. (1964).

Aitken, A. C., *Determinants and Matrices*, John Wiley & Sons, New York (1956).

Arfken, George B. and Weber, Hans J., *Mathematical Methods for Physicists*, Fourth Edition, Academic Press, San Diego (1995).

Bak, Thor A. and Lichtenberg, Jonas, *Mathematics for Scientists*, Benjamin Company, White Plains, New York (1964).

Bateman, H., *Partial Differential Equations of Mathematical Physics*, Dover Publications, New York (1944).

Boas, Mary L., *Mathematical Methods in the Physical Sciences*, John Wiley & Sons, New York (1983).

Borisenko, A. I. and Taropov, I. E., *Vector and Tensor Analysis with Applications*, Prentice-Hall, Englewood Cliffs, New Jersey (1968).

Brady, Wray G. and Mansfield, Maynard J., *Calculus*, Little, Brown and Company, Boston (1960).

Brand, Louis, *Vector Analysis*, John Wiley & Sons, New York (1957).

Brauer, Frec and Nohel, John A., *Differential Equations: A First Course* (second edition), Benjamin Company, Menlo Park, California (1973).

Braun, Martin, *Differential Equations and their Applications* (fourth edition), Springer-Verlag, New York (1993).

Brown, W. C., *Matrices and Vector Spaces*, Marcel Dekker, New York (1991).

Buck, R. C. and Buck, Ellen F., *Advanced Calculus* (third edition), McGraw-Hill Book Company, New York (1978).

Byron, F. W. Jr. and Fuller, R. W., *Mathematics of Classical and Quantum Physics*, Addison-Wesley Publishing Company, Reading, Massachusetts (1969).

Carslaw, H. S., *Introduction to the Theory of Fourier's Series and Integrals* (third edition), Dover Publications, New York (1952).

Champeney, D. C., *Fourier Transforms and Their Physical Applications*, Academic Press, New York (1973).

Churchill, R. V., Brown, J. W. and Verkey, R. F., *Complex Variables and Applications* (third edition), McGraw-Hill Book Company, New York (1974).

Cohen, H., *Mathematics for Scientists and Engineers*, Prentice-Hall, Englewood Cliffs, New Jersey (1992).

Courant, R. and Hilbert, D., *Methods of Mathematical Physics* (vol. I), John Wiley & Sons, New York (1953). Translated from the German.

Davis, H. F. and Snider, A. D., *Introduction to Vector Analysis* (fourth edition), Allyn & Bacon, Boston (1979).

Davis, P. J. and Rabinowitz, P., *Numerical Integration*, Blaisdell, Waltham, Massachusetts (1967).

Dwass, Meyer, *Probability: Theory and Applications*, Benjamin Company, White Plains, New York (1970).

Frazer, R. A., Duncan, W. J. and Collar, A. R., *Elementary Matrices and Some Applications in Dynamics and Differential Equations*, Cambridge University Press, Cambridge (1946).

Fuller, Gordon and Parker, Robert M., *Analytic Geometry and Calculus*, D. Van Nostrand Company, Princeton, New Jersey (1964).

Garcia, A. L., *Numerical Methods for Physics*, Prentice-Hall, Englewood Cliffs, New Jersey (1994).

Goldberg, Samuel, *Probability: An Introduction*, Prentice-Hall, Englewood Cliffs, New Jersey (1960).

Gradshteyn, I. S. and Ryzhik, I. M., *Table of Integrals, Series and Products*, Academic Press, New York (1980).

Hamming, R. W., *Numerical Methods for Scientists and Engineers* (second edition), McGraw-Hill Book Company, New York (1973).

Ince, E. L., *Ordinary Differential Equations*, Dover Publications, New York (1926).

Jeffreys, H. and Jeffreys, B. S., *Methods of Mathematical Physics* (third edition), Cambridge University Press, Cambridge (1966).

Kaplan, Wilfred, *Advanced Calculus*, (second edition), Addison-Wesley Publishing Company, Cambridge, Massachusetts (1973).

Lomont, J. S., *Applications of Finite Groups*, Academic Press, New York (1959).

Love, Clyde E. and Rainville, Earl D., *Differential and Integral Calculus* (fifth edition), The Macmillan Company, New York (1954).

Luke, Y. L., *Mathematical Functions and their Approximations*, Academic Press, New York (1975).

Maeder, Roman, *Programming in Mathematica* (third edition), Addison Wesley Publishing Company, Reading, Massachusetts (1997).

Margeneau, Henry and Murphy, George M., *The Mathematics of Physics and Chemistry* (second edition), D. Van Nostrand Company, Princeton, New Jersey (1964).

Miller, R. K. and Michel, A. N., *Ordinary Differential Equations*, Academic Press, New York (1982).

Morse, P. M. and Feshbach, H., *Methods of Theoretical Physics*, McGraw-Hill Book Company, New York (1953).

Parratt, Lyman G., *Probability and Experimental Errors in Science*, John Wiley & Sons, New York (1961).

Peirce, B. O., *A Short Table of Integrals* (third revised edition), Ginn and Company, Boston (1929).

Pipes, L. A. and Harvil, L. R., *Mathematics for Engineers and Physicists* (third edition), McGraw-Hill Book Company, New York (1970).

Press, W. H., Flannery, B. P., Teucholsky, S. A. and Vetterling, W. T., *Numerical Recipes* (second edition), Cambridge University Press, Cambridge (1992).

Rainville, E. D., *Infinite Series*, The Macmillan Company, New York (1967).

Rainville, E. D., *Special Functions*, The Macmillan Company, New York (1960).

Ralston, Anthony and Rabinowitz, Philip, *A First Course in Numerical Analysis*, McGraw-Hill Book Company, New York (1978).

Rubinstein, Isaak and Rubinstein, Lev, *Partial Differential Equations in Classical Mathematical Physics*, Cambridge University Press, Cambridge (1993).

Sneddon, I. N., *Special Functions of Mathematical Physics and Chemistry* (third edition), Longman Publishing Group, New York (1980).

Spiegel, M. R., *Theory and Problems of Complex Variables*, McGraw-Hill Book Company, New York (1964).

Spiegel, Murray R. *Schaum's Outline of Theory and Problems of Advanced Calculus*, McGraw-Hill Book Company, New York (1963).

Spiegel, Murray R., *Schaum's Outline of Theory and Problems of Complex Variables*, McGraw-Hill Book Company, New York (1964).

Spiegel, Murray R., *Schaum's Outline of Theory and Problems of Fourier Analysis*, McGraw-Hill Book Company, New York (1974).

Spiegel, Murray R., *Schaum's Outline of Theory and Problems of Laplace Transforms*, McGraw-Hill Book Company, New York (1965).

Spiegel, Murray R., *Schaum's Outline of Theory and Problems of Vector Analysis and an Introduction to Tensor Analysis*, McGraw-Hill Book Company, New York (1959).

Stoer, J. and Burlirsch, R., *Introduction to Numerical Analysis*, Springer-Verlag, New York (1980).

Tebbutt, Peter, *Basic Mathematics for Chemists*, John Wiley & Sons, Chichester (1994).

Thomas, G. B. Jr. and Finney, R. L., *Calculus and Analytic Geometry*, (seventh edition), Addison-Wesley Publishing Company, Reading, Massachusetts (1990).

Vance, Elbridge P., *Modern Algebra and Trigonometry*, Addison-Wesley Publishing Company, Reading, Massachusetts (1962).

Whittaker, E. T. and Watson, G. N., *A Course of Modern Analysis*, Cambridge University Press, Cambridge (1973).

Wrede, R. C., *Introduction to Vector and Tensor Analysis*, John Wiley & Sons, New York (1963). Dover Publications, New York (1972).

Young, Hugh D., *Statistical Treatment of Experimental Data*, McGraw-Hill Book Company, New York (1962).

Chemical physics

Allen, Harry C. Jr. and Cross, Paul C., *Molecular Vib-Rotors*, John Wiley & Sons, New York (1963).

Atkins, P. W., *Molecular Quantum Mechanics* (second edition), Oxford University Press, London (1983).

Atkins, P. W., *Quanta* (second edition), Oxford University Press, London (1991).

Bearman, Richard J. and Chu, Benjamin, *Problems in Chemical Thermodynamics*, Addison-Wesley Publishing Company, Reading, Massachusetts (1967).

Bellamy, L. J., *The Infrared Spectra of Complex Molecules*, vol. I, John Wiley & Sons, New York (1954); vol. II (3rd ed.), Chapman and Hall, New York (1980).

Brillouin, Léon, *Wave Propagation in Periodic Structures* (second edition), McGraw-Hill Book Company, New York (1946); Dover Publications, New York (1953).

Buerger, M. J., *Elementary Crystallography*, John Wiley & Sons, New York (1956).

Califano, S., *Vibrational States*, John Wiley & Sons, Chichester (1976).

Chamberlain, John, *The Principles of Interferometric Spectroscopy*, John Wiley & Sons, Chichester (1979).

Conners, Kenneth A., *Chemical Kinetics*, VCH Publishers, New York (1990).

Cotton, F. Albert, *Chemical Applications of Group Theory* (third edition), John Wiley & Sons, New York (1990).

Coulson, C. A., *Valence* (second edition), Oxford University Press, London (1961).

Decius, J. C. and Hexter, R. M., *Molecular Vibrations in Crystals*, McGraw-Hill Book Company, New York (1977).

Eyring, Henry, Walter, John and Kimball, George E., *Quantum Chemistry*, John Wiley & Sons, New York (1944).

Glasstone, Samuel, *Theoretical Chemistry*, D. Van Nostrand Company, New York (1944).

Golden, Sidney, *An Introduction to Theoretical Physical Chemistry*, Addison-Wesley Publishing Company, Reading, Massachusetts (1961).

Golden, Sidney, *Elements of the Theory of Gases*, Addison-Wesley Publishing Company, Reading, Massachusetts (1964).

Herzberg, G., *Molecular Spectra and Molecular Structure II. Infrared and Raman Spectra of Polyatomic Molecules*, D. Van Nostrand Company, Princeton, New Jersey (1945).

Hirschfelder, J. G., Curtis, C. F. and Bird, R. B., Molecular Theory of Gases and Liquids, John Wiley & Sons, New York, (1954).

Hollas, J. Michael, *Modern Spectroscopy* (second edition), John Wiley & Sons, Chichester (1992).

Jackson, J. D., *Classical Electrodynamics* (second edition), John Wiley & Sons, New York (1975).

Kemp, Merwin K., *Physical Chemistry. A Step-by-Step Approach*, Marcel Dekker, New York (1979).

Laidler, K. J., *Chemical Kinetics*, McGraw-Hill Book Company, New York (1950).

Levine, Ira N., *Molecular Spectroscopy*, John Wiley & Sons, New York (1975).

Lowe, John P., *Quantum Chemistry* (second edition), Academic Press, London (1993).

March, Norman H. and Mucci, Joseph F., *Chemical Physics of Free Molecules*, Plenum Press, New York (1993).

McQuarrie, Donald A., *Statistical Mechanics*, HarperCollins Publishers, New York (1976).

Messiah, A., *Quantum Mechanics* (vols. I & II), North-Holland Press, Amsterdam (1961). Translated from the French.

Mills, Ian *et al.* (eds), *Quantities, Units and Symbols in Physical Chemistry*, "The Green Book", Blackwell Scientific Publications, London (1993).

Milonni, Peter W. and Eberly, Joseph H., *Lasers*, John Wiley & Sons, New York (1988).

Moore, Walter J., *Physical Chemistry*, Prentice-Hall, Englewood Cliffs, New Jersey (1955).

Panofsky, W. K. H. and Phillips, M., *Classical Electricity and Magnetism* (second edition), Addison-Wesley, Reading, Massachusetts (1962).

Partington, J. R., *An Advanced Treatise on Physical Chemistry*, Longmans, Green and Company, London (1949).

Pauling, Linus and Wilson, E. Bright Jr., *Introduction to Quantum Mechanics*, McGraw-Hill Book Company, New York (1935). Republished by Dover Publications, Mineola, New York, (1985).

Pauling, Linus, *General Chemistry*, W. H. Freeman and Company, San Francisco (1947).

Rothschild, W. G., *Dynamics of Molecular Liquids*, John Wiley & Sons, New York (1984).

Schiff, Leonard I., *Quantum Mechanics*, McGraw-Hill Book Company, New York (1955).

Shubnikov, A. V. and Belov, N. V., *Colored Symmetry*, Pergamon Press, Oxford (1964).

Slater, John C., *Introduction to Chemical Physics*, McGraw-Hill Book Company, New York (1939).

Slater, John C., *Quantum Theory of Molecules and Solids* (vol. I), McGraw-Hill Book Company, New York (1963).

Sokolov, A. A., Loskutov, Y. M. and Ternov, I. M., *Quantum Mechanics*, Holt, Rinehart and Winston, New York (1966). Translated from the Russian.

Sverdlov, L. M., Kovner, M. A. and Krainov, E. P., *Vibrational Spectra of Polyatomic Molecules*, John Wiley & Sons, New York. Translated from the Russian.

Tinkham, Michael, *Group Theory and Quantum Mechanics*, McGraw-Hill Book Company, New York (1964).

Turrell, George, *Gas Dynamics. Theory and Applications*, John Wiley & Sons, Chichester (1997).

Turrell, George, *Infrared and Raman Spectra of Crystals*, Academic Press, London (1972).

Wigner, Eugene P., *Group Theory (and its application to the quantum mechanics of atomic spectra)*, Academic Press, New York (1959).

Wilson, E. Bright Jr., *An Introduction to Scientific Research*, McGraw-Hill Book Company, New York (1952).

Wilson, E. Bright Jr., Decius, J. C. and Cross, P. C., *Molecular Vibrations*, McGraw-Hill Book Company, New York (1955).

Zewail, Ahmed (ed.), *The Chemical Bond. Structure and Dynamics*, Academic Press, London (1992).

Author Index

Subject Index