

Orbitals in Chemistry

A Modern Guide for Students

VICTOR M. S. GIL

University of Coimbra, Portugal



CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK <http://www.cup.cam.ac.uk>
40 West 20th Street, New York, NY 10011-4211, USA <http://www.cup.org>
10 Stamford Road, Oakleigh, Melbourne 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain

© Victor M. S. Gil 2000

This book is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2000

Printed in the United Kingdom at the University Press, Cambridge

Typeface Times 11/14pt. System 3b2 6.03d [ADVENT]

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

Library of Congress Cataloguing in Publication data

Gil, Victor M. S., 1939–

Orbitals in chemistry: modern guide for students/Victor M. S. Gil.
p. cm.

Includes bibliographical references.

ISBN 0-521-66167-6 (hb)

1. Molecular orbitals. I. Title.

QD461 G52 2000

541.2'8–dc21 99-461968

ISBN 0 521 66167 6 hardback

ISBN 0 521 66649 X paperback

Contents

<i>Preface</i>	ix
<i>Acknowledgements</i>	xii
1. Energy, probability and electrons	1
1.1 Energy quantization	1
1.2 The wave–particle duality, observations and probability	6
1.3 Wavefunctions and the indeterminacy principle	10
2. An introduction to the dynamics of microsystems	18
2.1 Operators and observables	18
2.2 Expectation values of observables	21
2.3 Commuting operators	23
2.4 Important operators	24
2.5 The Schrödinger equation	27
2.6 A simple system: translational motion of a particle	30
2.7 Relativity theory, quantum mechanics and spin	42
3. One-electron atoms: atomic orbitals	45
3.1 Wave equation and angular momentum	45
3.2 Atomic orbitals	51
3.3 Spin	67
4. The one-electron molecule H_2^+: molecular orbitals	70
4.1 The wave equation and molecular orbitals	70
4.2 Molecular orbitals from atomic orbitals	74
4.3 Classifying molecular orbitals and electronic states	75

5. Many-electron atoms and the orbital concept	85
5.1 Wavefunction and the Pauli principle	85
5.2 Electron repulsion: orbitals, an approximation	90
5.3 Total electronic energy	95
5.4 Orbital energies	97
5.5 Electronic configurations	101
5.6 Beyond electronic configurations: terms, levels, states	105
5.7 Density-functional theory and Kohn–Sham orbitals	111
5.8 Relativistic corrections	112
6. Orbitals in diatomic molecules	114
6.1 The approximations	114
6.2 The simple diatomics H_2 , He_2^+ and ‘ He_2 ’	115
6.3 Molecular orbitals in X_2 molecules	120
6.4 Heterodiatomic molecules	129
6.5 Electronegativity	136
7. Orbitals in polyatomic molecules	139
7.1 New features relative to diatomic molecules	139
7.2 Molecular orbitals in AH_n molecules	140
7.3 Other molecules and quantitative m.o. theory	156
8. Molecular orbitals and electron pair bonding	167
8.1 Atoms in molecules and structural formulae	167
8.2 The theory of atoms in molecules	169
8.3 Structural formulae and non-independent bonds	175
8.4 Orbitals and electron pairing in valence-bond theory	180
8.5 Molecular geometry and the valence-shell electron pair repulsion model	185
8.6 Canonical molecular orbitals and localized functions	193
8.7 Use and misuse of the hybrid orbital concept	201
9. π Molecular orbitals: conjugation and resonance	205
9.1 The σ – π separation	205
9.2 The CO_2 molecule and the CO_3^{2-} ion	206
9.3 The ethylene and acetylene molecules	214
9.4 The butadiene molecule	217
9.5 The benzene molecule	223
9.6 π Electron densities and bond orders	227

10. Patterns in localized chemical bonds	230
10.1 Back to structural formulae	230
10.2 Bond energies and the Periodic Table	232
10.3 The octet rule and the writing of structural formulae	236
10.4 The conservation of the sum of bond orders	242
11. The concept of molecular orbitals in other systems	245
11.1 The C ₆₀ molecule	245
11.2 Octahedral complexes of transition metals	248
11.3 The band theory of solids	258
12. Orbitals in action	265
12.1 Orbitals and chemical reactivity	265
12.2 Orbitals and spectroscopy	273
<i>Answers to problems</i>	290
<i>References</i>	303
<i>Index</i>	311

1

Energy, probability and electrons

1.1 Energy quantization

The interpretation of chemical phenomena, and hence the development of chemistry, owes a great deal to two fundamental concepts: *energy* and *probability*. The scientific idea of energy only emerged during the nineteenth century, whereas the notion of probability is at least two centuries older. In modern chemistry, there is hardly any field which does not depend upon one or both of these basic concepts, quite often coupled to each other. Important examples where energy and probability converge simultaneously are the second law of thermodynamics and entropy (in the field of chemical transformations), and quantum chemistry and orbitals (in the field of structure of atoms and their groupings). Orbitals are the main subject of this book. In so far as the orbital concept is essential to the study of the structure of matter it also lies in the realm of physical properties – such as electric, magnetic, spectroscopic and related properties – and of chemical behaviour of substances, from both the point of view of kinetics and thermodynamics. As we will briefly see below, the links between the fields from which the two previous examples were drawn have long been established; especially the historical role played by thermodynamics in the foundations of our knowledge about matter and light is recalled.

Orbitals are mathematical functions of the coordinates of each electron in atoms, molecules and other atomic aggregates; we will see that they carry the dimensions $[\text{length}]^{-3/2}$. They contain information on the *probability distribution* for each electron in space and correspond to certain electronic *energy values*. There are two most significant non-classical features which concern the energy of bound electrons. One is that not all values of energy are allowed: *energy quantization*. The other is the so-called *exchange energy* which is a consequence of the indistinguishability of electrons. In addition,

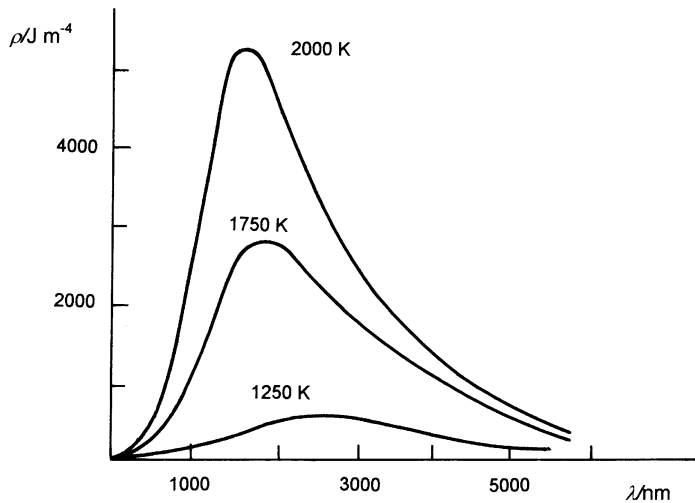


Fig. 1.1 Distribution of the energy output for a black-body (the ideal case of a body that neither reflects nor transmits light; it just absorbs or emits) at different temperatures as a function of wavelength.

our knowledge of the position and momentum of electrons is affected by the *Heisenberg indeterminacy principle*.

The concept of energy quantization has its roots in two kinds of experimental data gathered during the second half of the nineteenth century, both related to light: (a) the discontinuous emission spectra of gaseous elements; and (b) the distribution of the light intensity emitted by heated bodies as a function of wavelength for various temperatures (the so-called black-body) (Fig. 1.1).

The discrete wavelengths λ of the emission spectra were found by the end of the nineteenth century to be reproduced by a simple empirical formula, the Balmer–Rydberg equation:

$$1/\lambda = \mathcal{R}(1/n_1^2 - 1/n_2^2) \quad (1.1)$$

where $\mathcal{R} = 1.097 \times 10^5\ \text{cm}^{-1}$ is a constant (the Rydberg constant) and n_1 and n_2 ($n_1 < n_2$) are integer numbers. At about the same time (1890), the German physicist Max Planck (1858–1947, 1918 Nobel laureate in Physics) found another empirical formula which reproduces the density of radiating energy due to a black-body, as a function of frequency (ν) and temperature (T):

$$\rho(\nu, T) = A\nu^3/(e^{B\nu/T} - 1) \quad (1.2)$$

where A and B are constants.

The strange results summarized in Eq. (1.1) were in apparent conflict with thermodynamic data, especially with the measured heat capacities of gaseous elements. Meanwhile, Planck, taking advantage of his experience in thermodynamics, sought an explanation of Eq. (1.2) on the basis of the statistical interpretation of entropy that had been developed in 1896 by another Austrian physicist, Ludwig E. Boltzmann (1844–1906). Following previous ideas, Planck began by considering the atoms at the surface of a black-body as oscillators capable of absorbing and emitting light. In his studies and as a mathematical convenience, Boltzmann had often begun by considering the energy in small discrete amounts and only at the end of his analysis allowed this discontinuity to be removed. Planck used a similar treatment and considered discrete values $0, h\nu, 2h\nu, 3h\nu, \dots$ (more or less populated according to the temperature) for the energies of the atomic oscillators, ν being the frequency of the oscillation and h a constant that, at the final stage, should be set equal to zero to allow all vibrational energies to be taken into account. It was then found by Planck that agreement with experiment (that is, Eq. (1.2)) could only be reached if h remained non-zero and was set equal to 6.63×10^{-34} J s. The theoretically derived counterpart of Eq. (1.2) is

$$\rho(\nu, T) = (8\pi h\nu^3/c^3)/(e^{h\nu/kT} - 1) \quad (1.3)$$

where c is the speed of light in vacuum and k is the Boltzmann constant. The constant h was later, very properly, named the Planck's constant. Meanwhile, $h\nu$, the difference between successive energy values, was called a *quantum* of energy for the electromagnetic oscillators. The concept of energy quantization was born (with important contributions from the work of other physicists, mainly Franck and Hertz). Although the energy difference $h\nu$ between successive energy values was later found to be in agreement with the quantum-mechanical treatment of diatomic molecules as harmonic oscillators, the minimum value, zero, considered by Planck should be replaced by $h\nu/2$, the so-called zero-point energy. For an harmonic oscillator, the allowed vibrational energy values are $h\nu/2, 3h\nu/2, 5h\nu/2, \dots$

In 1905, Albert Einstein (1879–1955, 1921 Nobel laureate in Physics), born in Germany and later a naturalized American, having overcome some initial disbelief concerning the energy quantization of oscillators, extended this idea to electromagnetic radiation itself. Radiation carries energy in discrete amounts – *quanta* – each quantum or ‘particle’ of radiating energy being

$$E = h\nu \quad (1.4)$$

where ν is now the frequency of the electromagnetic radiation. It has been recognized that Einstein became much more enthusiastic about Planck's theory than the author himself, who had great difficulty in discarding the principles of mechanics – now called classical mechanics – on which he had been brought up. The 'particles' of radiating energy were some years later named *photons* by the American chemist Gilbert N. Lewis (1875–1946). Thus, the energy of η photons corresponding to a given frequency ν is

$$E = \eta h\nu. \quad (1.5)$$

The photon model explained more than the black-body radiation. One of the first great achievements of the new theory of light was the interpretation of the *photoelectric effect*. The difference between the energy $h\nu$ of the incident photon and the minimum energy I necessary to remove an electron from the structure it belongs to (ionization energy, in the case of gaseous samples; work function, in the case of condensed phases) appears as kinetic energy $mv^2/2$ of the ejected electron:

$$h\nu = I + mv^2/2. \quad (1.6)$$

Thus, electron ejection requires a minimum frequency ν , it being irrelevant to have an intense source of light (many photons per unit of area) if the frequency is less than that minimum. However, the more intense the beam of light of appropriate frequency the greater the number of electrons ejected, by a one-to-one photon–electron interaction.

The proposal of Einstein was not accepted easily. Even as late as 1913, Planck himself, when joining other distinguished German physicists in recommending Einstein's appointment to the Prussian Academy of Sciences, would write

... That he may sometimes have missed the target in his speculations, as, for example, in his hypothesis of light quanta, cannot really be held against him, for it is not possible to introduce fundamentally new ideas, even in the most exact sciences, without occasionally taking a risk.

The idea of energy quantization was brought into chemistry with the application of quantum theory to the electronic structure of atoms in 1913 by the Danish physicist Niels Bohr (1885–1962, 1922 Nobel laureate in Physics). At the time, Bohr was working in the laboratory of the New Zealand physicist Ernest Rutherford (1871–1937, 1909 Nobel laureate in Chemistry) in England, a short time after the nuclear structure for the atom had been established by Rutherford and his co-workers. Classical electromagnetic theory predicted that the electrons around the nucleus,

because they are being accelerated (centripetal acceleration), should radiate energy and thus should gradually approach the nucleus, emitting energy in a continuous manner, until the final collapse into the nucleus. However, it was known that atoms are not only stable entities but also, when previously excited, they radiate energy in discrete amounts. Then, inspired by the quantum theory of Planck, Bohr was forced to admit that the electrons in atoms exist in *stationary energy states*, with a well-defined energy, which they can only leave by either absorption or emission of certain discrete values of energy.

In such states, the electrons (charge $-e$ and mass m_e) would have circular orbits (radius r) around the nucleus, undergoing transitions between orbits through either absorption or emission of energy. In particular, for the hydrogen atom, the total energy of the electron in circular orbit (the kinetic energy in the rest frame of the nucleus, $m_e v^2/2$, plus the potential energy associated with the electron–nucleus attraction, $-\hbar e^2/r$) could only have certain values. In order to obtain these values, Bohr began by assuming, as was done by Planck for the radiation of the black-body, that the energy emitted by the excited H atom was given by

$$E = n' h \nu' \quad (1.7)$$

with n' a positive integer and ν' a frequency characteristic of the motion of the electron around the nucleus. The crucial problem was then to relate ν' with the angular velocity of the electron circular motion. Bohr made some conjectures that, however, led to an impasse. When, almost by chance, the Balmer–Rydberg equation (1.1) was made known to him, the solution became clear and the following formula could be established for the possible values of energy of the electron in the H atom:

$$E_n = -(2\pi^2 e^4 m_e) / h^2 n^2 \quad (1.8)$$

with $n = 1, 2, 3, \dots$ a *quantum number*.

This energy quantization would imply, too, quantization of the angular momentum of the electron $m_e v r$:

$$m_e v r = n h / 2\pi. \quad (1.9)$$

Contrary to what is often found in books, Eq. (1.9) was not the starting point for the energy quantization expression (1.8), but the other way round. (For discussion of this point and the presentation of the Bohr model, see, for example, refs. 1 and 2.)

However, the theory presented several limitations and could not explain the spectra of atoms other than mono-electronic atoms. Bohr himself would recognize the need for a new theory and, indeed, would contribute to it. In particular, it is noted that, although Eq. (1.8) is reproduced by quantum mechanics, the angular momentum of the electron is not given by Eq. (1.9); in particular, the minimum value possible is zero and not $h/2\pi$ as required by Eq. (1.9).

1.2 The wave–particle duality, observations and probability

The photon theory of light and Eq. (1.5) received additional confirmation when, in 1923, the American physicist Arthur H. Compton (1892–1962, 1927 Nobel laureate in Physics) discovered the effect that would bear his name: *the Compton effect*. When X-rays interact with electrons, the scattered radiation, after transferring some energy to stationary electrons (which are accelerated by the electric field of the radiation), has a slightly higher wavelength. In contrast with classical physics, this increase $\Delta\lambda$ depends only on the angle θ through which the radiation is deflected, being independent of the initial wavelength:

$$\Delta\lambda = \lambda' - \lambda = \lambda_c(1 - \cos\theta). \quad (1.10)$$

The constant $\lambda_c = 2.425$ pm is called the *Compton wavelength* of the electron. The wavelength shift given by Eq. (1.10) can be easily reproduced theoretically if the interaction between the radiation and the electron is considered as a collision between two particles in which the energy and the linear momentum are conserved (conservation of momentum in the incident direction and in the direction perpendicular to it). These particles are a photon of energy $h\nu$ and linear momentum $p = h\nu/c = h/\lambda$ and a stationary electron of mass m_e which acquires velocity v (Fig. 1.2). It is then found

$$\lambda_c = h/m_e c = 2.425 \text{ pm}. \quad (1.11)$$

Equation (1.5) establishes a bridge between a description of light as an (electromagnetic) wave of frequency ν and as a beam of η energy particles. If phenomena related to time averages, such as diffraction and interference, can be easily interpreted in terms of waves, other phenomena, involving a one-to-one relation such as the photoelectric and the Compton effects, require a description based on corpuscular attributes. This wave–particle duality reflects the use of one or the other description depending on the experiment performed, while no experiment exists which exhibits both aspects of the duality simultaneously.

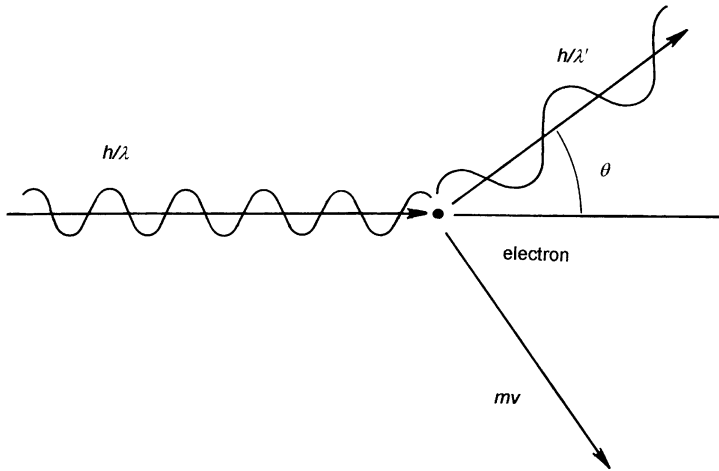


Fig. 1.2 Conservation of linear momentum in the Compton effect.

The form $p = h\nu/c = h/\lambda$ for the linear momentum of a photon can be obtained by combining the expression for the energy of a photon $E = h\nu$ and the expression $E = m_{\text{ph}}c^2$ which defines the relativistic mass of the photon, m_{ph} , provided that the linear momentum of a photon is made equal to $m_{\text{ph}}c$ by analogy with the classical expression mv for the linear momentum of a particle.

Strongly influenced by the interpretation of the Compton effect, the French physicist Louis Victor, Prince de Broglie (1892–1987, 1929 Nobel laureate in Physics), suggested in his doctoral thesis in 1924 that the wave–particle duality for photons could be extended to any particle of momentum $p = mv$ which, *somehow*, would then have a wavelength – the *de Broglie wavelength* – associated with it and given by

$$\lambda = h/p. \quad (1.12)$$

Indeed, considering that photons are rather peculiar particles in that they have zero rest mass and can exist only when travelling at the speed of light, it seems reasonable that the wave associated with the motion of any particle should become more and more apparent as the mass decreases, rather than the wave coming into existence suddenly when the rest mass vanishes.

This was a revolutionary and quite nebulous suggestion, not easily accepted at the time. If it was not for the intervention of Einstein, excited with the proposal, Louis de Broglie would most likely have failed his doctoral examination. Louis de Broglie found support for his hypothesis and attempted to clarify the ‘wave characteristics’ of a moving electron by

successfully reproducing the Bohr expression (1.9) for the quantization of the angular momentum of the electron in the H atom. By trying to adjust waves to a circular orbit, it is found that destructive interference between waves in successive cycles is only avoided if the orbit perimeter is a whole number of wavelengths: $2\pi r = n\lambda$. Associating Eq. (1.12), and putting $p = m_e v$, Eq. (1.9) is then reproduced.

The ‘associated wave’ of the de Broglie relation was later replaced by the quantum-mechanical wavefunction of the particle. But the experimental proof of the de Broglie hypothesis and his quantitative relation appeared in 1927 with the first observations of diffraction of electrons by C. Davisson and L. Germer in America and by G.P. Thomson in Great Britain, using a nickel crystal and a gold foil, respectively. These results were found to be analogous to those obtained when using X-rays having wavelengths equal to the de Broglie wavelengths for electrons. It is interesting to note that, whereas J.J. Thomson showed at the end of the nineteenth century that the electron ‘is a particle’, and received the 1906 Nobel prize in Physics mainly for that, his son G.P. Thomson showed that it ‘is a wave’, and got the same prize in 1937 (shared with Davisson). In 1932, the occurrence of diffraction of helium atoms and hydrogen molecules by crystals was also found to be in agreement with the de Broglie relation.

The interpretation given by de Broglie for the quantization of the angular momentum of the electron of H, in the Bohr model, assumes *in some sense* that the electron can interfere with itself. Although any diffraction experiment always presupposes a large number of particles, it is not necessary that all the particles are considered at the same time; diffraction can still be obtained with a sequence of a large number of particles, one at a time. For example, a diffraction pattern is gradually built up as photons or electrons pass through two slits which are separated by a distance of the order of the wavelength of the particles (Fig. 1.3). In this sense, it can be said that each particle interferes with itself (refs. 3–5).



Fig. 1.3 Gradual genesis of an electron interference pattern in a double-slit experiment, with electrons reaching the detector one at a time. (Adapted with permission from ref. 5.)

This result clearly points to the impossibility of considering that the particle passes through one slit and not through the other; it seems ‘to pass through both’. In other words, we can only say that it has 50% probability of passing through one slit and 50% of passing through the other. A bridge between the *wave characteristics* of small particles and *probability* is beginning to arise. Simultaneously, the notion of trajectory – from one slit to the screen or detector – no longer applies.

In contrast to this, for macroscopic particles, it is possible to identify the slit through which they pass; trajectories can then be defined and no interference occurs. The passages through each slit are then independent events and the probability of particles striking a given point on the screen, in a certain time interval, is just the sum of the probabilities corresponding to passage through slit 1 and through slit 2.

The interference of microscopic particles leads to a diffraction pattern with deviations with respect to the mere sum of the individual probabilities. The two events are no longer independent. If we wish to state in advance where the next particle will appear, we are unable to do so. The best we can do is to say that the next particle is more likely to strike in one area than another. *A limit to our knowledge*, associated with the wave–matter duality, becomes apparent. In the double-slit experiment, we may know the momentum of each particle but we do not know anything about the way the particles traverse the slits. Alternatively, we could think of an experiment that would enable us to decide through which slit the particle has passed, but then the experiment would be substantially different and the particles would arrive at the screen with different distributions. In particular, the two slits would become distinguishable and independent events would occur. No interference would be detected, that is, the wave nature of the particle would be absent. In such an experiment, in order to obtain information about the particle position just beyond the slits, we would change its momentum in an unknown way. Indeed, recent experiments have shown that interference can be made to disappear and reappear in a ‘quantum eraser’ (ref. 6 and references therein).

This discussion extends to electrons, and small particles in general, what we have already found for photons, that is, the appearance of electrons behaving as classical particles or as a wave depends on the experiment being performed. On the other hand, it illustrates a basic feature of microsystems: there is an unavoidable and uncontrollable interaction between the observer and the observed. No matter how cleverly one devises an experiment, there is always some disturbance involved in any measurement and such a disturbance is intrinsically indeterminate (see also Section 1.3).

However, the situation concerning the limits to our knowledge depicted above is not as desperate as it would seem. The fact that we can say that the next particle in a double-slit diffraction experiment is more likely to strike in one area of the screen than another points to the existence of a certain ‘determinism’ in the statistical result for a large number of identical experiments. We need to consider not each individual microsystem but a *statistical ensemble*, that is a great number of non-interacting replicas of a given microsystem which enable a large number of identical experiments to be performed; or, alternatively, we need to repeat the same experiment with the same microsystem, always in the same conditions. For example, one mole of H atoms corresponding to 6.022×10^{23} atoms all in the same electronic state is an ensemble for measurements about the electron. This characterization of a collection of microsystems in the same conditions is called *state preparation*.

Each individual measurement of any physical quantity yields a value A . But, independently of any possible observation errors associated with imperfect experimental measurements, the outcomes of identical measurements in identically prepared microsystems are not necessarily the same. The results fluctuate around a central value. It is this collection or spectrum of values that characterizes the *observable* A for the ensemble. The fraction of the total number of microsystems leading to a given A value yields the probability of another identical measurement producing that result. Two parameters can be defined: the *mean value* (later to be called the ‘expected value’) and the *indeterminacy* (also called uncertainty by some authors). The mean value $\langle A \rangle$ is the weighted average of the different results considering the frequency of their occurrence. The indeterminacy ΔA is the standard deviation of the observable, which is defined as the square root of the dispersion. In turn, the dispersion of the results is the mean value of the squared deviations with respect to the mean $\langle A \rangle$. Thus,

$$\Delta A = \langle (A - \langle A \rangle)^2 \rangle^{1/2} \quad (1.13)$$

which is a statistically meaningful expression of precision or reproducibility of measurements. (For a discussion of precision and accuracy in measurements, see, for example, ref. 7.)

1.3 Wavefunctions and the indeterminacy principle

Much of what has been said earlier lies at the foundations of the new mechanics which began with the work of the German physicist Max Born

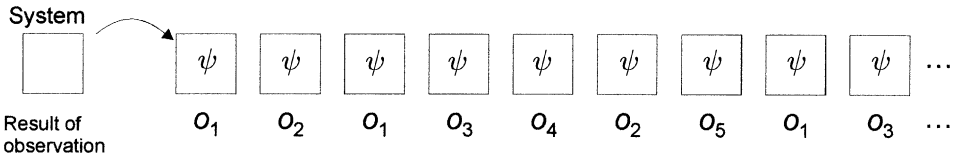


Fig. 1.4 A quantum ensemble is constructed by replicating a microsystem many times, all members prepared in the same way so that they are in the same state (represented by ψ).

(1882–1970, 1954 Nobel laureate in Physics) in 1924. In quantum mechanics, two ensembles which show the same distributions for all the observables are said to be in the same *state*. Although this notion is being introduced for statistical ensembles, it can also be applied to each individual microsystem (see, for example, ref. 8), because all the members of the ensemble are identical, non-interacting and identically prepared (Fig. 1.4). Each state is described by a *state function*, ψ (see, for example, ref. 3). This state function should contain the *information* about the probability of each outcome of the measurement of any observable of the ensemble. The wave nature of matter, for example the interference phenomena observed with small particles, requires that such state functions can be superposed just like ordinary waves. Thus, they are also called *wavefunctions* and act as probability amplitude functions.

In particle diffraction experiments, the probability distribution $P(x, t)$ of the particles, electrons say, in the detector (x coordinate) should be reproduced from a wavefunction $\psi(x, t)$. At each value of time, the function $P(x, t)$ is large at values of x where an electron is likely to be found and small at values of x where an electron is unlikely to be found. Using $\psi(x, t)$ as a probability amplitude function which, to be more general, is allowed to be a complex function, the relation is

$$P(x, t) \propto \psi^*(x, t)\psi(x, t) \quad (1.14)$$

where $\psi^*(x, t)$ is the complex conjugate of $\psi(x, t)$: if $\psi(x, t)$ is of the form $a + ib$, $\psi^*(x, t)$ is $a - ib$ and $P(x, t) = a^2 + b^2$, a necessarily positive quantity. The relation (1.14) resembles that between the probability of finding photons at a given point and the square of the electromagnetic field at that point.

In a double-slit experiment, two functions can be defined, $\psi_1(x, t)$ and $\psi_2(x, t)$, for the two possibilities of the electron behaviour, one corresponding to passage through slit 1, the other to slit 2. The appropriate wavefunction

is the *superposition* of $\psi_1(x, t)$ and $\psi_2(x, t)$, that is,

$$\psi(x, t) = \psi_1(x, t) + \psi_2(x, t), \quad (1.15)$$

and the probability density function of the electrons at the detector is then given by

$$\begin{aligned} \psi^*(x, t)\psi(x, t) &= [\psi_1^*(x, t) + \psi_2^*(x, t)][\psi_1(x, t) + \psi_2(x, t)] \\ &= \psi_1^*(x, t)\psi_1(x, t) + \psi_2^*(x, t)\psi_2(x, t) + \psi_1^*(x, t)\psi_2(x, t) \\ &\quad + \psi_2^*(x, t)\psi_1(x, t). \end{aligned} \quad (1.16)$$

The last two terms in Eq. (1.16) represent interference. This can be made more clear if the complex functions are written in terms of their magnitudes and phases:

$$\begin{aligned} \psi_1(x, t) &= |\psi_1(x, t)| \exp[i\alpha_1(x, t)] \\ \psi_2(x, t) &= |\psi_2(x, t)| \exp[i\alpha_2(x, t)]. \end{aligned} \quad (1.17)$$

It is easily found that substitution of Eq. (1.17) into Eq. (1.16) leads to:

$$\psi^*(x, t)\psi(x, t) = |\psi_1(x, t)|^2 + |\psi_2(x, t)|^2 + 2|\psi_1(x, t)| |\psi_2(x, t)| \cos(\alpha_1 - \alpha_2). \quad (1.18)$$

The interference term depends on the phase difference $\alpha_1 - \alpha_2$ of the probability amplitude functions.

We have mentioned in Section 1.2 that when enough information is available to determine which path is taken by the particle, the interference disappears, and it reappears when the path cannot be determined. The explanation for this is not that quantum waves refrain from interfering when observed too closely, it is just that information about the path is available when the quantities measured are not sensitive to the phase difference $\alpha_1 - \alpha_2$, but quantities that are sensitive to this phase difference can be measured when those that determine the path are not (ref. 6).

The example of the double-slit experiment already shows that the physical information contained in a state function is inherently probabilistic in nature. In the next chapter this feature will be further developed leading to the central concept of this book: *orbitals*. For example, the orbitals of the H atom are wavefunctions $\psi(x, y, z)$ that enable the electron probability density to be known: $\psi^*(x, y, z) \psi(x, y, z)$. The idea of a trajectory (or orbit) is replaced by the idea of a probability distribution.

The above interpretation of the wavefunction, linking a mathematical function with a physical state, is due to Born and provides the basis of the most widely held view of quantum mechanics: the *Copenhagen interpretation*. Helped by this interpretation, other scientists rapidly developed the theoretical machinery for extracting information about position, momentum and other observables from wavefunctions, during the years 1925–1927. In 1925, the German physicist Werner Heisenberg (1901–1976, 1932 Nobel laureate in Physics), disciple of Max Born, represented the physical quantities by complex numbers in a mathematical treatment that Born clarified by establishing a relation with the properties of matrices. This work was in part developed in collaboration with the German physicist Pascual Jordan whom Born met, by accident, during a train journey.

In the same year, the English mathematician (of Swiss father) Paul Dirac (1902–1984, 1933 Nobel laureate in Physics) developed a new formalism strongly related to classical mechanics through a collection of postulates, which is equivalent to that of Heisenberg but more easy to understand. Almost at the same time, the Austrian physicist Erwin Schrödinger (1887–1961, 1933 Nobel laureate in Physics, prize shared with Dirac) proposed a wave approach to quantum mechanics, inspired by the de Broglie theory which he had initially classified as nonsense. As we shall see in Chapter 2, the form of his wave equation is analogous to the equations of classical electromagnetic theory, the wavelength being given by the de Broglie relation. It is a differential equation, involving second-order derivatives of the state function of the system, ψ , one of the unknowns, and a second unknown which is the energy E . By introducing basic conditions which ψ must obey because of its relation with probability densities, the solution of the wave equation can lead to a discrete set of E values; once these values are known, the wavefunctions ψ associated with them are then obtained. In this way, energy quantization appears as a direct consequence of the physical interpretation of the wavefunction.

Initially, Schrödinger's theory was not well accepted by either Heisenberg or Dirac, but gained the support of Born. The situation changed when, in 1926, Schrödinger showed that his theory was equivalent to those of Heisenberg and Dirac. Beginning with the Heisenberg–Born equations involving matrices to represent physical quantities, Schrödinger showed that these could be represented by appropriate *operators* (see Chapter 2). In the particular case of the electron in the H atom, the operator corresponding to the observable energy is related to the Hamilton function of classical mechanics, modified previously by Dirac.

In 1927, Heisenberg identified *incompatible observables* whenever the commutative property of multiplication does not apply to the corresponding

operators and showed that any attempt to simultaneously measure two incompatible physical quantities A and B necessarily introduces an imprecision in each observable. No matter how the experiment is designed, the results are inevitably indeterminate, and the indeterminacies ΔA and ΔB cannot be reduced to zero. Instead their product is always larger than a constant of the order of Planck's constant. The easiest practical way to recognize incompatible variables is to note that their dimensions multiply to joule second. For example, position x (y , or z) and momentum p_x (p_y , or p_z) are fundamentally incompatible observables, in the sense that knowing the precise value of one precludes knowing anything about the other. Thus,

$$\Delta x \Delta p_x \geq h/4\pi \quad \Delta y \Delta p_y \geq h/4\pi \quad \Delta z \Delta p_z \geq h/4\pi. \quad (1.19)$$

This is known as the *Heisenberg indeterminacy principle* (also called uncertainty principle by some authors). It has to do with precision and not with accuracy. This situation has already been met in Section 1.2 when referring to the double-slit experiment.

On the scale of classical mechanics, the indeterminacies can, at least in principle, be made sufficiently small to be negligible. Then, the order in which we measure position and linear momentum is arbitrary. Irrespective of this order, we may have, for example, 2 m for the position x of a body (with respect to some origin) and 3 kg m s^{-1} for its momentum p_x which implies $x p_x = 2 \text{ m} \times 3 \text{ kg m s}^{-1} = p_x x = 3 \text{ kg m s}^{-1} \times 2 \text{ m} = 6 \text{ kg m}^2 \text{ s}^{-1} = 6 \text{ J s}$. However, for an ensemble of microsystems, the indeterminacies are no longer simultaneously negligible. In such a case, we need matrices or operators to represent those observables, and $x p_x \neq p_x x$. Heisenberg showed that $x p_x - p_x x = i\hbar/2\pi$ (see Chapter 2). This is equivalent to saying that the order of preparation of the ensemble for adequate measurements is not irrelevant, as shown below.

Let us assume that we begin by preparing a state ψ so that identical measurements of the momentum p_x in the various members of the ensemble have an arbitrarily small dispersion. Then, if, in a separate experiment we prepare the same state ψ and make identical individual measurements of the position x , the result is a huge dispersion of the x values. If, on the other hand, we begin by forcing (preparing) the ensemble to a state ψ' in order to yield an arbitrarily small indeterminacy of position, then there will be a huge indeterminacy in the momentum values obtained through identical measurements of the various microsystems in state ψ' (Fig. 1.5). The impossibility of reducing both indeterminacies to an arbitrarily small value is intimately related to the fact that ψ is different from ψ' . In Chapter 2, we will say that no 'eigenfunction' for the position operator (corresponding to a

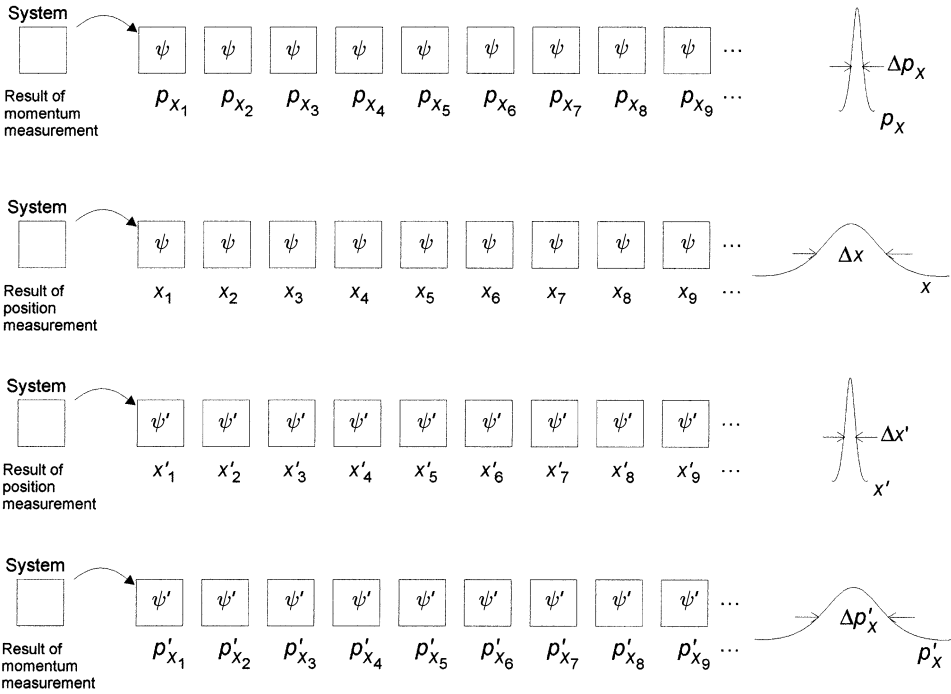


Fig. 1.5 Identical preparation of an ensemble for measurements of two incompatible observables A and B , so that one of them is affected by an arbitrarily small indeterminacy; the indeterminacy in the other is necessarily huge.

dispersion-free distribution of results) is an eigenfunction for the momentum operator, and vice versa.

It should be noted that, since we are talking about measurements in separate experiments, strictly speaking it is not the measurement of x (or p_x) that directly affects the measurement of p_x (or x); it is the preparation process of the ensemble for the two measurements which, being the same, relates the two measurements. Thus, the usual statement *the impossibility of knowing simultaneously with accuracy (rather, precision) both position and momentum of a particle* should be read as *the impossibility of preparing a state for which both position and momentum can be determined with arbitrarily small indeterminacies* (see, for example, ref. 9).

The limitation reflected in the Heisenberg principle is *implicit in Nature*. It has nothing to do with a particular apparatus or with the imperfections of experimental techniques. A particle *cannot* simultaneously have a precise value of x and a precise value of p_x . This is different from assuming that each particle really does have definite values for both position and momentum,

but these definite values cannot be determined because measurement of one property alters the value of the other (see, for example, ref. 10).

What is being said applies to incompatible observables. For example, the position in coordinate x and the linear momentum in the y direction are not incompatible. Therefore, it is possible to have an ensemble prepared in such a way as to enable arbitrarily small indeterminacies in both x and p_y . That is, it is possible to find identical wavefunctions irrespective of the order of the preparations for both measurements ($\psi = \psi'$) and

$$\Delta x \Delta p_y \leq h/4\pi. \quad (1.20)$$

The connection with wave-matter duality can be further developed (see, for example, ref. 11). For an ensemble in a state of well-defined linear momentum p (along the x direction, say) we can define a unique de Broglie wavelength $\lambda = h/p$ and the particle can be at any point (in the x direction) with equal probability (this point will be considered again in Chapter 2). Conversely, if we know the position of the particle precisely, then the wavefunction must assume a large value at the point considered and very small values outside. Since such a function can be obtained by superposing a large number of waves of various wavelengths which interfere constructively at the point and destructively otherwise, the existence of various λ values means that all information about the linear momentum is lost (Fig. 1.6).

Obviously, if we cannot specify the position and the momentum of a particle, at any instant, we cannot maintain the concept of trajectory for small particles. This idea involves statements about the past, the present and the future of the system, and only the present is accessible to experimental determination and, even so, with limitations. According to Heisenberg, it

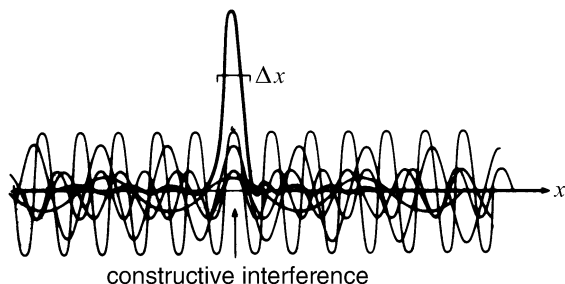


Fig. 1.6 A small indeterminacy in x as the result of appropriate superposition of waves having a variety of wavelengths: information about linear momentum $p = h/\lambda$ is entirely lost (see, for example, ref. 12).

makes no sense to ascribe properties to a system that cannot, at least in principle, be subjected to experimental determination. Properties of macro-systems exist by themselves, independently of being measured, and then we can speak both of precision and accuracy of measurements. This is not the case with microsystems. Here, properties are latent until they are the object of observation and measurement (at least in principle), and precision rules.

To conclude this introduction to quantum mechanics, it is interesting to note the omnipresence and the agglutinating role of Planck's constant. Indeed, if it was set equal to zero, all the construction which began with black-body radiation and the quantization of radiation energy, followed by the wave-matter duality and the Heisenberg principle ... would fall down. In addition, the intrinsic angular momentum (spin) of some particles, including the electron, would be forced to be zero, with many consequences at the theoretical and practical levels.