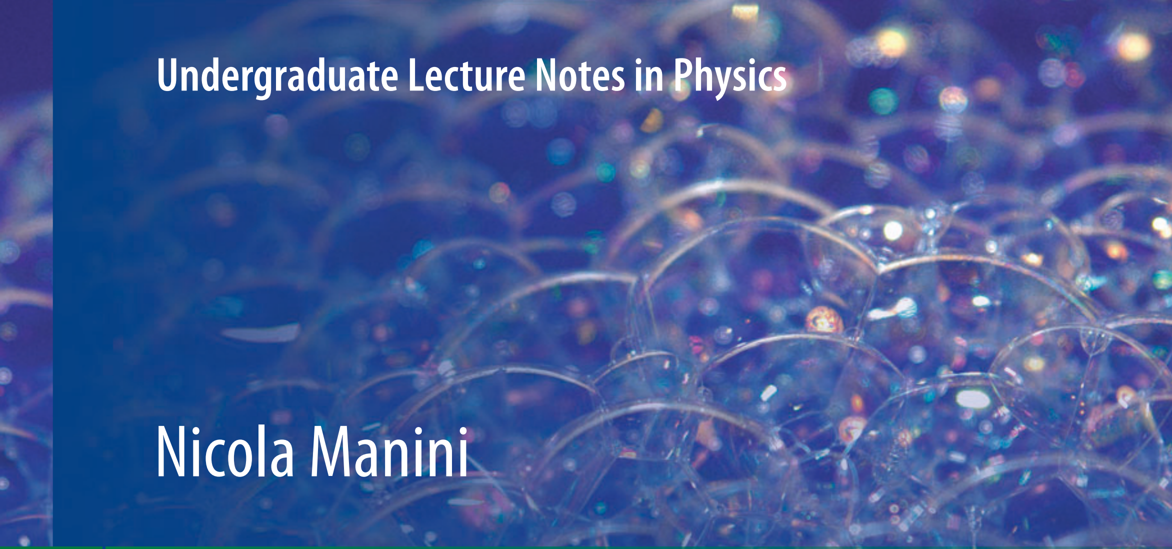
Undergraduate Lecture Notes in Physics



Nicola Manini

Introduction to the Physics of Matter

Basic Atomic, Molecular, and Solid-State Physics

Springer

Undergraduate Lecture Notes in Physics

Undergraduate Lecture Notes in Physics (ULNP) publishes authoritative texts covering topics throughout pure and applied physics. Each title in the series is suitable as a basis for undergraduate instruction, typically containing practice problems, worked examples, chapter summaries, and suggestions for further reading.

ULNP titles must provide at least one of the following:

- An exceptionally clear and concise treatment of a standard undergraduate subject.
- A solid undergraduate-level introduction to a graduate, advanced, or nonstandard subject.
- A novel perspective or an unusual approach to teaching a subject.

ULNP especially encourages new, original, and idiosyncratic approaches to physics teaching at the undergraduate level.

The purpose of ULNP is to provide intriguing, absorbing books that will continue to be the reader's preferred reference throughout their academic career.

More information about this series at http://www.springer.com/series/8917

Nicola Manini

Introduction to the Physics of Matter

Basic Atomic, Molecular, and Solid-State Physics

Springer

Nicola Manini
Department of Physics
University of Milan
Milan
Italy

ISSN 2192-4791 ISSN 2192-4805 (electronic)
Undergraduate Lecture Notes in Physics
ISBN 978-3-319-14381-1 ISBN 978-3-319-14382-8 (eBook)
DOI 10.1007/978-3-319-14382-8

Library of Congress Control Number: 2014958570

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

This book fulfills a twofold purpose: to provide a pedagogic panorama of the current microscopic understanding of the basic physics of matter, and to help students acquire a quantitative feeling of the typical orders of magnitude of the main physical quantities (energy, time, temperature, length) involved in the specific conditions relevant for "matter" in its atomic, molecular, and condensed forms. Both tasks are favored by keeping structurally and conceptually well distinct the analysis of the adiabatically separate motions of electrons and of atoms. This distinct treatment is organized in close parallel for molecules and for solids.

While keeping different degrees of freedom well distinct, formal likeness is noted whenever useful, following the standard strategy of *similar solutions for similar equations*. Noteworthy examples of this approach include the spherically symmetric motion of electrons in atoms and of nuclei in diatomic molecules, as well as applications of the Fermi-gas model to electrons in metals and to fluid ^{3}He.

Several detailed derivations are included, when pedagogically useful to understand the physical reasons of some results. Most of these derivations deserve being understood and then quickly forgotten. In contrast, a number of mathematical relations summarize essential physical information and results, and are therefore well worth retaining. As a guide for the reader, a gray background highlights these *essential equations*.

Numbers and orders of magnitude are important. At least as much as mathematical derivations, probably more. A broad selection of numerical problems invites the reader to familiarize herself with the conceptually simple but often practically intricate numerical evaluations and unit conversions required to reach quantitatively correct estimates in real life or laboratory conditions.

At variance with many textbooks in this field, the present one adopts SI units throughout. This choice does not only follow international recommendation, but also helps comparing all the results with the output of instruments. The one indulgence to non-SI units is a frequent quotation of energies in eV, which represent a practical unit for most atomic-scale phenomena, and converts easily to Joule.

As a basic textbook, this one focuses on what is currently well under control: a selection of well-understood systems, phenomena, experimental techniques, and conceptual schemes. From this panorama, the reader should be warned against gathering the false impression that the last word has been said about the physics of matter. On the contrary, physicists, chemists, and biologists currently investigate matter in its multifaceted forms, and collect a wealth of experimental data, for which understanding is often only qualitative and partial. Creativity and insight help scientists to develop novel approximate schemes (models) to interpret these data and proceed toward a better understanding of the intimate structure and dynamics of matter.

Not everything is equally important. A reader wishing to focus on a bare minimum contents, while still grasping the essentials of the physics of matter is suggested to skip the following nonessential topics:

- the spectral-broadening mechanisms—Sect. 1.2,
- the hyperfine structure of H—Sect. 2.1.8,
- perturbation theory applied to the 2-electron atom—Sect. 2.2.3,
- the details of the variational calculation of Sect. 3.2.1,
- the electronic molecular transitions discussed in Sect. 3.3,
- the density operator formalism—Sect. 4.1.2,
- the foundation of the canonical ensemble—Sect. 4.2 before Eq. (4.10),
- the connection of entropy and statistics—Sect. 4.2.2,
- the laser—Sect. 4.4.1,
- the tight-binding and plane waves models for the electron bands of crystals—Sect. 5.2.1,
- the extrinsic semiconductors and their technological applications, from page 214 to the end of Sect. 5.2.2.

On the other hand, a reader who aims at a broader view of the field than is provided by the topics selection covered by the present textbook can follow the extended introductory treatments, e.g., in Refs. [1–4]. Specialized texts focus on advanced approaches closer to the frontiers of research, covering both the theoretical and the experimental side. The reader is invited to browse in particular Refs. [5–7] for atomic physics, Refs. [6–9] for molecules, and Refs. [10–12] for solid-state physics. Finally, this textbook focuses on the present-day understanding of the physics of matter, omitting most of the fundamental experiments and conceptual steps through which the scientists reached this understanding. Reference [13] provides an insider's view into the historical evolution of the basic concepts in this field.

The present volume draws its initial inspiration from Luciano Reatto's course delivered in the 1990s at the University of Milano. In the past, the bulk of this volume was made available as lecture notes, first released on January 15, 2004. Precious feedback and suggestions from colleague Giovanni Onida and from many students in the Milan Physics Diploma course prompted numberless corrections, refinements, and clarifications through over one decade. Compared to the previous lecture notes, the present Springer edition benefits also from (i) the addition of an Appendix summarizing the minimal quantum mechanics needed to understand the physics of matter, (ii) the addition of selected problems, and (iii) a thorough revision of the text and figures. The problems included in this book, plus many more, are available at the website http://materia.fisica.unimi.it/manini/dida/archive_exam.html.

The author acknowledges the warm feedback and interest from students of physics at the University of Milano: this book was originally written for them. Last but not least, the patient care and love by the author's family was a primary ingredient in nurturing the present textbook.

Milano, November 2014 Nicola Manini

Contents

Chapter 1
Introductory Concepts

1.1 Basic Ingredients

A substantial body of experimental evidence accumulated mainly through the late 19th and early 20th century eventually convinced the community of physicists and chemists that any piece of matter (e.g. pure helium gas in a vessel, a block of solid ice, a metallic screw, a mobile phone, a block of wood, a bee…) is ultimately composed of a (often huge) number of electrons and atomic nuclei.

An electron is a *bona-fide* elementary point-like particle characterized by a mass $= m_e \simeq 0.911 \times 10^{-30}$ kg, and a negative charge $= -q_e$, where the elementary charge $q_e \simeq 1.60 \times 10^{-19}$ C.

Nuclei come with a complicated inner structure, involving length scales $\simeq 10^{-15}$ m and excitation energies $\simeq 10^{-13}$ J. These nuclear properties are largely irrelevant to the "ordinary" properties of matter: to most practical purposes one can describe matter by modeling nuclei as approximately structureless point-like particles. A nucleus containing Z protons and A nucleons (protons and neutrons alike), has positive charge $= Zq_e$ and mass $M \simeq A\,\text{a.m.u.} = A \times 1.66 \times 10^{-27}\,\text{kg} \gg m_e$.

If we neglect relativistic effects and the interaction of a piece of matter with its surroundings, then all internal microscopic interactions among its components are of simple electromagnetic nature. The non-relativistic motion of nuclei and electrons in the sample is governed by the following (Hamiltonian) energy operator:

$$H_{\text{tot}} = T_n + T_e + V_{ne} + V_{nn} + V_{ee} \tag{1.1}$$

where:

$$T_n = \frac{1}{2}\sum_{\alpha} \frac{P_{R_\alpha}^2}{M_\alpha} \tag{1.2}$$

N. Manini, *Introduction to the Physics of Matter*,
Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-14382-8_1

is the kinetic energy of the nuclei (P_{R_α} is the conjugate momentum to the position R_α),

$$T_e = \frac{1}{2m_e} \sum_i {P_{r_i}}^2 \tag{1.3}$$

is the kinetic energy of the electrons (P_{r_i} is the conjugate momentum to r_i),

$$V_{ne} = -\frac{q_e^2}{4\pi\varepsilon_0} \sum_\alpha \sum_i \frac{Z_\alpha}{|R_\alpha - r_i|} \tag{1.4}$$

is the potential energy describing the attraction between nuclei and electrons,

$$V_{nn} = \frac{q_e^2}{4\pi\varepsilon_0} \frac{1}{2} \sum_\alpha \sum_{\beta\neq\alpha} \frac{Z_\alpha Z_\beta}{|R_\alpha - R_\beta|} \tag{1.5}$$

describes the nucleus–nucleus repulsion, and finally

$$V_{ee} = \frac{q_e^2}{4\pi\varepsilon_0} \frac{1}{2} \sum_i \sum_{j\neq i} \frac{1}{|r_i - r_j|} \tag{1.6}$$

represents the electron–electron repulsion. Basically, the distinction between a steel key and a bottle of beer is the result of their different "ingredients", i.e. the number of electrons and the number and types of nuclei (charge numbers Z_α and masses M_α) involved.

A state ket $|\xi\rangle$ containing all quantum-mechanical information describing the motion of all nuclei and electrons evolves according to Schrödinger's equation

$$i\hbar \frac{d}{dt} |\xi(t)\rangle = H_{\text{tot}} |\xi(t)\rangle. \tag{1.7}$$

This equation, based on Hamiltonian (1.1), is apparently simple and universal. This simplicity and universality indicates that in principle it is possible to understand the observable behavior of any isolated macroscopic object in terms of its microscopic interactions. In practice, however, exact solutions of Eq. (1.7) are available for few simple and idealized cases only. If one attempts an approximate numerical solution of Eq. (1.7), (s)he soon faces the problem that the information contents of a N-particles ket increases exponentially with N, and soon exceeds the capacity of any computer. To describe even a relatively basic material such as a pure rarefied molecular gas, or an elemental solid, nontrivial approximations to the solution of Eq. (1.7) are called for.

The application of smart approximations to Eq. (1.7) to understand observed properties and to predict new properties of material systems is a refined art. These approximations often represent important conceptual tools linking the macroscopic properties of matter to the underlying microscopic interactions. The present textbook proposes a panoramic view of several observed phenomena in the physics of matter, introducing a few standard conceptual tools for their understanding. The proposed schemes of approximation represent a pedagogical selection of rather primitive idealizations: the bibliography at the end suggests directions to expand the reader's conceptual toolbox to approach today's state of the art in research. We should be aware of the limitations of state-of-the-art tools: even smart and experienced physicists of matter risk to deliver inaccurate predictions of a basic property such as the electrical conductivity of a pure material of known composition and structure, before actually measuring that property. For more complex systems (e.g. biological matter), quantitative and often even qualitative predictions based on Eq. (1.7) outrun the capability of today's modeling capability and computing power.

1.1.1 Typical Scales

The motions described by Hamiltonian $H_{\rm tot}$ involve several characteristic dimensional scales, dictated by the physical constants [14] in $H_{\rm tot}$, where the absence of the speed of light c is noteworthy. Firstly, observe that in $H_{\rm tot}$ the elementary charge q_e and electromagnetic constant ε_0 always appear in the fixed combination

$$e^2 \equiv \frac{q_e^2}{4\pi\varepsilon_0} = 2.3071 \times 10^{-28}\,\mathrm{J\,m},$$

of dimensions energy × length. A unique combination of e^2, Planck's constant $\hbar$, and electron mass m_e yields the characteristic length

$$a_0 = \frac{\hbar^2}{m_e e^2} = 0.529177 \times 10^{-10}\,\mathrm{m} \qquad (1.8)$$

named *Bohr radius*, which sets the typical length scale of electronic motions. Most microscopic structures and patterns of matter arise naturally with spacings of the order of a_0. For example, atomic positions can be probed by means of scanning microscopes. These instruments slide a very sharp tip over a solid surface: the atomic force microscope (AFM) maps the forces that the surface atoms exert as they come into contact with the tip; the scanning tunneling microscope (STM) maps an electronic tunneling current between the tip and the surface as they are kept a fraction of nm apart. This (and other) class of experiments provide consistent evidence, e.g. Fig. 1.1, that indeed in materials atoms are typically spaced by a fraction of nm, namely approximately $2 \div 10$ times a_0.

The interaction energy of two elementary point charges at the typical distance a_0

$$E_{\rm Ha} = \frac{e^2}{a_0} = \frac{m_e e^4}{\hbar^2} = 4.35975 \times 10^{-18}\,\mathrm{J} = 27.2114\,\mathrm{eV}, \qquad (1.9)$$

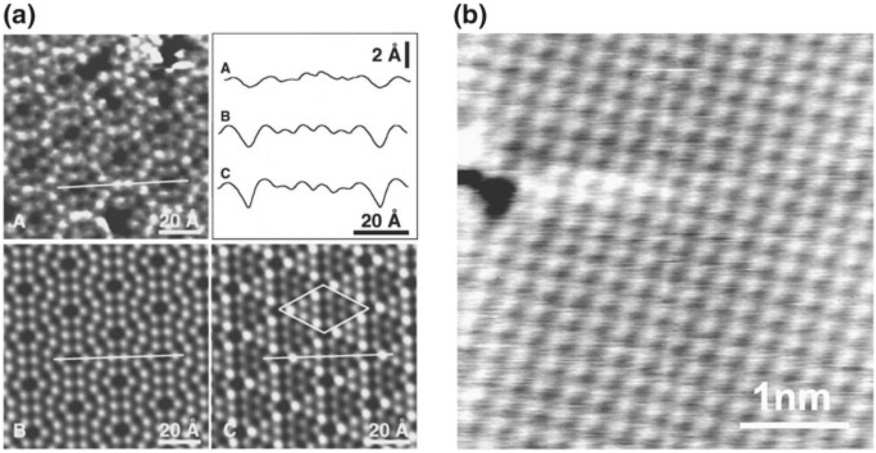


Fig. 1.1 **a** The AFM (*A*) and STM (*B, C*) topography images of a clean silicon (111) surface. The STM map visualizes empty (*B*) and filled (*C*) electronic states. Reprinted figure with permission from R. Erlandsson, L. Olsson, and P. Mårtensson, Phys. Rev. B **54**, R8309 (1996). Copyright (1996) by the American Physical Society. **b** The AFM topography of a clean copper (100) surface. Reprinted figure with permission from Ch. Loppacher, M. Bammerlin, M. Guggisberg, S. Schär, R. Bennewitz, A. Baratoff, E. Meyer, and H.-J. Güntherodt, Phys. Rev. B **62**, 16944 (2000). Copyright (2000) by the American Physical Society

named *Hartree energy*, sets a natural energy scale for phenomena involving one electron in ordinary matter. In practice, the eV ($\simeq 0.037\, E_{\rm Ha}$) is a more commonly used energy unit. The nuclear charge factors $Z_\alpha \lesssim 10^2$ can scale the e^2 electron–nucleus coupling constant up by $\lesssim 10^2$, the electron–nucleus distance down by $\gtrsim 10^{-2}$, and therefore increase the binding energies of electrons by up to 4 orders of magnitude ($10^4\, E_{\rm Ha} \simeq 300\,$keV). On the other hand, delicate balances may occasionally yield electronic excitation energies as small as 1 meV. The motions of the nuclei are usually associated to smaller energies ($\sim 10^{-4} \div 10^{-3}\, E_{\rm Ha}$) and velocities than electronic motions, because of the at least 1,836 times larger mass at the denominator of the kinetic term of Eq. (1.2).

The typical timescale of electronic motions is inversely proportional to its energy scale:

$$t_0 = \frac{\hbar}{E_{\rm Ha}} = \frac{\hbar^3}{m_e e^4} = 2.41888 \times 10^{-17}\ \text{s}. \tag{1.10}$$

Oscillations of period $2\pi t_0$ have a frequency $\nu_0 = \omega_0/(2\pi) = E_{\rm Ha}/(2\pi\hbar) = 6.5797 \times 10^{15}\,$Hz.

The typical electron velocity is then set by the ratio

$$v_0 = \frac{a_0}{t_0} = \sqrt{\frac{E_{\rm Ha}}{m_e}} = \frac{e^2}{\hbar} = 2.18769 \times 10^6\ \text{m/s}, \tag{1.11}$$

about 1 % of the speed of light, which justifies *a posteriori* the initial neglect of relativity. Specifically, the ratio

$$\alpha = \frac{v_0}{c} = \sqrt{\frac{E_{\mathrm{Ha}}}{m_e c^2}} = \frac{e^2}{\hbar c} = 7.29735 \times 10^{-3} = \frac{1}{137.036}, \qquad (1.12)$$

called *fine-structure constant*, measures the relative importance of the relativistic corrections to the nonrelativistic motion of the electrons. Heavier nuclei usually move far more slowly than electrons, thus relativistic corrections to their motion are usually negligible.

Figure 1.2 compares the typical length and frequency scales of electrons in matter to the wavelengths and frequencies of the electromagnetic waves, used to investigate matter itself as sketched mainly in Sects. 1.2, 4.4, 5.1.3, and 5.2.3. The radiation wavelength matches the typical interatomic distances ($\sim 10^{-10}$ m) in the X-rays region. On the other hand, radiation frequencies (and thus radiation energy quanta, or *photons*, see Sect. 4.3.2.2) match the typical atomic frequency scale ν_0 (and thus energy scale E_{Ha}) in the ultraviolet (UV) region: in this spectral region, the typical electromagnetic wavelength is a fraction of µm, i.e. three orders of magnitude larger than the typical atomic size $\sim a_0$. Typical frequencies ($\sim$5 THz) and energies ($\sim$10 meV) associated to the motion of the nuclei in matter belong to the infrared (IR) region of the electromagnetic spectrum.

1.1.2 Structure of Matter: A Panoramic View

In ordinary matter, electrons, driven by the attraction, Eq. (1.4), tend to lump around the strongest positive charges around, namely the nuclei. Despite the diverging attraction, due to the kinetic energy of Eq. (1.3) and Heisenberg's uncertainty relation (see Sect. B.2.1), electrons fail to collapse inside the nuclei, and rather form *atoms* of finite size $\approx a_0$, as discussed in Chap. 2. Atoms then act as the building blocks of matter, in its gaseous and condensed phases. Many observed macroscopic properties of extended matter such as elasticity, heat transport, heat capacity can be described in terms of the motion of atoms or aggregations of atoms. Ultimately, the interaction among atoms, governing these motions, is driven by the quantum dynamics of charged electrons and nuclei described by Eqs. (1.1) and (1.7), usually described within the *adiabatic separation* scheme (Sect. 3.1). Understanding the dynamics of a finite number of electrons in the field of two or several nuclei, and the motion of these nuclei themselves (Chap. 3) provides the basics of interatomic *bonding*, the mechanism granting the very existence of *molecules* and of condensed states of matter. The methods and approximations developed for few-atom systems and for large statistical ensembles (Chap. 4) lead eventually to new concepts and phenomena associated to the macroscopic size of extended objects (Chap. 5).

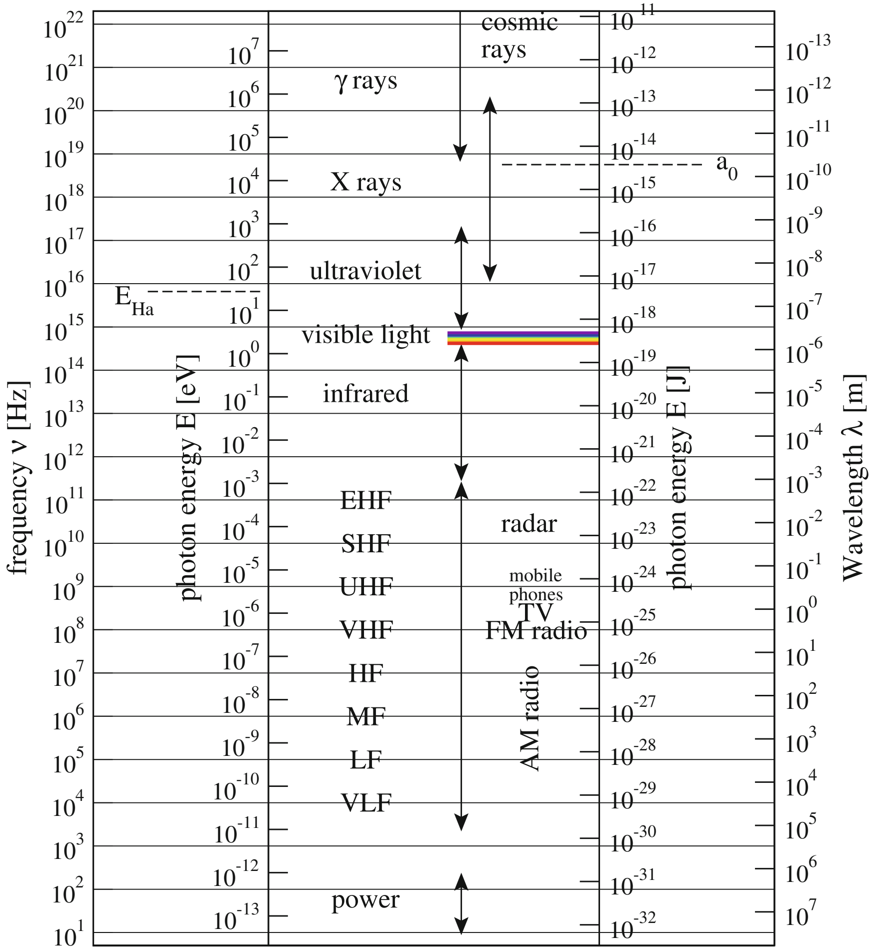


Fig. 1.2 The spectrum of electromagnetic radiation, characterized by frequency ν, photon energy $\mathscr{E} = \hbar\omega = 2\pi\hbar\nu$, and wavelength $\lambda = c/\nu$. Note the colors of visible light and the characteristic atomic quantities a_0, Eq. (1.8), and E_{Ha}, Eq. (1.9)

1.2 Spectra and Broadening

Starting from the late 19th century, physicists have developed and employed all sorts of techniques to investigate the intimate excitations of matter. A rich body of evidence has been collected through two broad classes of *spectroscopies*: *absorption* and *emission*.

- *Absorption*: a collimated beam of monochromatic light (not necessarily visible) crosses the sample; if the light frequency matches a characteristic sample transition, the photons are absorbed. The energy removed from the original beam

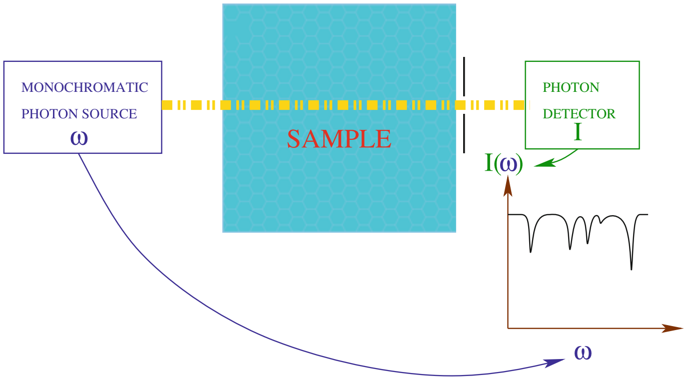
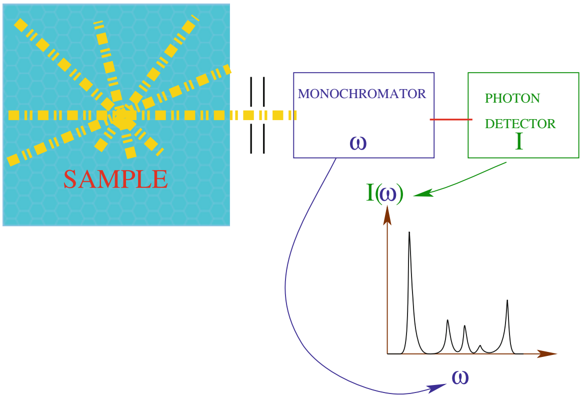


Fig. 1.3 A conceptual scheme of the setup for absorption spectroscopy

Fig. 1.4 A conceptual scheme of the setup for emission spectroscopy



is transformed, e.g. re-emitted in random directions (mostly different from the original one): the intensity loss of the beam as it crosses the sample is recorded as a function of light frequency (or equivalently wavelength), as sketched in Fig. 1.3.[1]

- *Emission*: the sample is brought to excited state(s), e.g. by means of a flame or by bombarding it with high-energy electrons. The light that the sample subsequently emits in the deexcitation transitions is collected, and its intensity is recorded as a function of frequency (Fig. 1.4).

Scientists record routinely spectra of both kinds in the IR, visible, UV and X-ray ranges, to probe the properties of gaseous (both atomic and molecular), liquid, and solid samples. Even the color analysis that the human eyes carry out for the light diffused by an object illuminated by white light can be qualified as a rough emission-spectroscopy experiment.

---

[1] The transmitted intensity $I(\omega)$ can be converted into an absorption coefficient $\alpha(\omega)$ by inverting the Beer-Lambert attenuation law $I/I_0 = \exp(-\alpha\, l\, N/V)$. Here I_0 is the original radiation intensity, l is the probed sample thickness, and N/V is the number density of absorbing atoms/molecules.

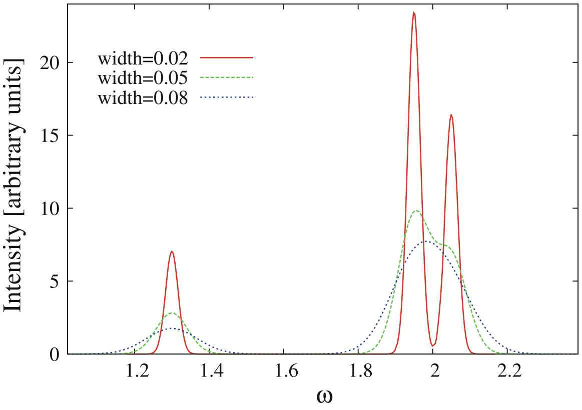


Fig. 1.5 Broadening limits the detail observable in a spectrum: when the line width exceeds the separation between two lines (*dotted curve*), they look like a single bump, to the extent that the "fine" detail of the two near lines is lost

Atomic and molecular spectra are often characterized by sharp monochromatic peaks (also called "lines") associated to transitions between quantum states $|i\rangle$ and $|f\rangle$, characterized by discrete energy levels E_i and E_f. The angular frequency of one such sharp peak is $\omega_{if} = |E_i - E_f|/\hbar$.

In practice, no absorption/emission peak is ever infinitely sharp. Broadening limits the spectral details which can be resolved, see Fig. 1.5. At least 3 simultaneous effects combine to broaden each spectral line: (i) the finite *experimental resolution* of the spectrometer, (ii) a "*natural*" *broadening* due to finite lifetime, (iii) the *Doppler broadening*.

Experimental resolution is typically limited by the monochromator resolving power, noise in the photon detector and electronics, inhomogeneities of the sample or of some external field applied to it. Resolution broadening is usually generated by several random concurrent effects, and determines therefore a Gaussian line shape. It has no fundamental nature, thus it can be (and often is) reduced by means of technical improvements.

Elementary Schrödinger theory predicts all atomic eigenstates to be stationary states, because this theory neglects the interaction of matter with the always present fluctuating radiation field. In reality, the ground state is the one really stationary state, while all excited states decay spontaneously to lower-energy eigenstates, due to that interaction. An individual excited atom decays at an unpredictable random future time. However a large number N_0 of initially excited atoms decays predictably: after a time t, an average number of atoms

$$[N](t) = N_0\, e^{-t\gamma} = N_0\, e^{-t/\tau}, \tag{1.13}$$

have not yet decayed; this decay law defines the single-atom *lifetime* τ and decay rate $\gamma = 1/\tau$ of that atomic level. This decay rate γ is constant in time (an atom becomes no more or less likely to decay at the next instant because it has already spent some time in the excited level) and defines a uniform random decay.[2] In practice, very few atoms last in an excited state longer than a few times its characteristic τ. Accordingly, τ sets the typical duration of any spectroscopy experiment involving that excited state. Due to the time-energy uncertainty, the energy of an atomic level cannot be measured with better precision than

$$\Delta E \simeq \frac{\hbar}{\tau} = \hbar \gamma. \tag{1.14}$$

The effect of the finite lifetime of atomic levels on the spectral lines is therefore a broadening of the otherwise infinitely sharp line. This "natural" broadening due to lifetime appears in the spectrum as a Lorentzian peak profile

$$I(\omega) = I_0 \, \frac{\gamma^2}{(\omega - \omega_{if})^2 + \gamma^2}, \tag{1.15}$$

around the original line position ω_{if}. Atomic excited states are characterized by typical lifetimes of several ns, thus by natural spectral broadening $\hbar\gamma$ of a fraction of μeV.

The random thermal motion in a gas-phase sample introduces an extra source of line broadening: the Doppler effect. When viewed from the lab frame, the intrinsic transition frequencies of atoms/molecules moving toward or away from the detector at a velocity v_x are blue or red shifted relative to those at rest. Since the thermal (molecular center-mass) velocities are non-relativistic, the Doppler frequency shift is given by the simple form

$$\omega = \omega_{if} \left(1 \pm \frac{v_x}{c}\right) \qquad \omega_{if} = \text{angular frequency at rest.} \tag{1.16}$$

The molecular velocities are random distributed [see Eq. (4.50) in Sect. 4.3.1] depending on the gas temperature T: the average number of molecules with velocity component v_x in the direction of detection is

$$\frac{dn(v_x)}{dv_x} = N \sqrt{\frac{M}{2\pi k_\mathrm{B} T}} \, \exp\left(-\frac{M\, v_x^2}{2k_\mathrm{B} T}\right), \tag{1.17}$$

---

[2] The total decay rate γ is the sum of the rates of decay to all lower-lying states. For example, as described in Sect. 2.1.9, the 3p state of H decays at a rate $\gamma_{3\mathrm{p}\to 1\mathrm{s}} = 1.67 \times 10^8\,\mathrm{s}^{-1}$ to the ground state, and at a rate $\gamma_{3\mathrm{p}\to 2\mathrm{s}} = 2.25 \times 10^7\,\mathrm{s}^{-1}$ to the 2s state (decay to state 2p is dipole-forbidden, thus occurs at a negligible rate). Therefore the 3p state empties at a total rate $\gamma_{3\mathrm{p}} = \gamma_{3\mathrm{p}\to 1\mathrm{s}} + \gamma_{3\mathrm{p}\to 2\mathrm{s}} = 1.90 \times 10^8\,\mathrm{s}^{-1}$.

where M is the molecular mass. Radiation intensity then spreads around the rest frequency ω_{if} as

$$I(\omega) = I_0 \exp\left[-\frac{Mc^2}{2k_B T}\left(\frac{\omega - \omega_{if}}{\omega_{if}}\right)^2\right]. \tag{1.18}$$

This represents a Gaussian broadening of full width at half-maximum

$$\Delta\omega_{\text{Doppler}} = \omega_{if}\sqrt{8\ln(2)\,\frac{k_B T}{Mc^2}}. \tag{1.19}$$

As $\lambda = 2\pi c/\omega$, the relative broadening in terms of wavelength is the same as for frequency:

$$\frac{\Delta\lambda_{\text{Doppler}}}{\lambda_{if}} = \sqrt{8\ln(2)\,\frac{k_B T}{Mc^2}}. \tag{1.20}$$

Due to the $M^{-1/2}$ dependence of the broadening width in Eqs. (1.19) and (1.20), for a given temperature lighter atoms/molecules move faster and are then more severely affected by Doppler broadening. The spectrum of atomic hydrogen is especially Doppler-broadened. For example, the H_α line, introduced below, of 300 K gas-phase H suffers of $\Delta\lambda_{\text{Doppler}}/\lambda_{if} \simeq 3 \times 10^{-6}$ (an absolute width $\hbar\Delta\omega_{\text{Doppler}} \simeq 10\,\mu\text{eV}$) broadening.

According to the laws of statistics, the individual sources of broadening combine into a total broadening $\Delta\omega_{\text{tot}} = \left[(\Delta\omega_{\text{exp}})^2 + \gamma^2 + (\Delta\omega_{\text{Doppler}})^2\right]^{1/2}$.

Chapter 2
Atoms

The importance of the spectroscopy of atoms and ions for the understanding of the whole physics of matter cannot be overestimated. The study of atoms starts off naturally from the exact dynamics of a single electron in the central field of one positively charged nucleus (Sect. 2.1) because, beside the intrinsic interest of this system, the notation and concepts developed here are at the basis of the language of all atomic physics. This language is then used to introduce (Sect. 2.2) the spectroscopy of many-electron atoms, clusters of 2 to about 10^2 electrons repelling each other, but trapped together in the attractive central field of a single nucleus.

2.1 One-Electron Atom/Ions

The one-electron atom is one of the few quantum problems whose Schrödinger equation (1.7) is solved exactly. Here, comparison of theory and experiment allows physicists to evaluate the limits of validity and predictive power of the quantum mechanical model Eqs. (1.1)–(1.6). When relativistic effects are included (Sect. 2.1.7), this model is found in almost perfect agreement with extremely accurate experimental data, all the tiny discrepancies being satisfactorily accounted for by a perturbative treatment of residual interactions (Sect. 2.1.8).

The solution of the Schrödinger equation (1.7), or rather its time-independent counterpart (B.30), for the one-electron atom is a basic exercise in quantum mechanics (QM). Both the V_{ee} and V_{nn} terms in Eq. (1.1) vanish, and only the nuclear and electronic kinetic energies plus the Coulomb attraction V_{ne} are relevant. Many textbooks [5, 15, 16] provide detailed solutions of the one-electron atom problem. Here we recall the general strategy and main results.

- **Separation of the center-of-mass motion**. In analogy to the solution of the classical Kepler-Newton two-body "planetary" problem, the position operator of the nucleus $\mathbf{R}$, of mass M, and that of the electron $\mathbf{r}_e$ are replaced by the combinations

$$\mathbf{R}_{\rm cm} = \frac{M\mathbf{R} + m_e\mathbf{r}_e}{M + m_e} \quad \text{and} \quad \mathbf{r} = \mathbf{r}_e - \mathbf{R}\,. \tag{2.1}$$

N. Manini, *Introduction to the Physics of Matter*,
Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-14382-8_2

In these new coordinates, the Hamiltonian separates into a purely kinetic term for the free translational motion of the atomic center of mass $\mathbf{R}_{\mathrm{cm}}$

$$H_{\mathrm{cm}} = -\frac{\hbar^2}{2(M+m_e)}\nabla^2_{\mathbf{R}_{\mathrm{cm}}}, \tag{2.2}$$

plus a Coulombic Hamiltonian for the relative coordinate $\mathbf{r}$

$$H_{\mathrm{Coul}} = -\frac{\hbar^2}{2\mu}\nabla^2_{\mathbf{r}} - \frac{Ze^2}{|\mathbf{r}|}, \tag{2.3}$$

where

$$\mu = \frac{Mm_e}{M+m_e} \tag{2.4}$$

is the *reduced mass* of this 2-particle system. The free global $\mathbf{R}_{\mathrm{cm}}$ translational motion is described trivially in terms of plane waves. The internal atomic dynamics is that of a single particle of mass μ in the same Coulombic central field as the original nucleus-electron attraction.

- **Separation in spherical coordinates**. To exploit the spherical symmetry of the potential, the Schrödinger equation is conveniently rewritten in polar coordinates r, θ, φ. By factorizing the total wavefunction $\psi(r,\theta,\varphi) = R(r)\Theta(\theta)\Phi(\varphi)$, the variables separate, and the original three-dimensional (3D) equation splits into three independent second-order equations for the r, θ, and φ motions:

$$-\frac{d^2\Phi}{d\varphi^2} = \eta\,\Phi, \tag{2.5}$$

$$-\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\,\frac{d\Theta}{d\theta}\right) + \frac{\eta}{\sin^2\theta}\,\Theta = \lambda\,\Theta, \tag{2.6}$$

$$-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \left[U(r) + \frac{\hbar^2}{2\mu}\frac{\lambda}{r^2}\right]R = ER. \tag{2.7}$$

Here we put a general function $U(r)$ in place of the potential energy $-Ze^2/r$, to emphasize that this same technique can be applied to any central potential (e.g. in Sect. 3.3).

- **Solution** of the separate eigenvalue problems: the differential equations are solved under the appropriate boundary conditions for $R(r)$, $\Theta(\theta)$ and $\Phi(\varphi)$. The eigenvalues η, λ, and E can assume only certain values, compatible with the boundary conditions[1]:

$$\eta = \eta_{m_l} = m_l^2, \qquad m_l = 0,\ \pm 1,\ \pm 2, \ldots \tag{2.8}$$

$$\lambda = \lambda_l = l(l+1), \qquad l\ = |m_l|,\ |m_l|+1,\ |m_l|+2, \ldots \tag{2.9}$$

---

[1] The "quantization" of the angular motion originates from the boundary conditions $\Phi(\varphi+2\pi) = \Phi(\varphi)$, and finite $\Theta(0)$ and $\Theta(\pi)$ for the solutions of Eqs. (2.5) and (2.6): these conditions are granted only for integer values of the quantum numbers m_l and $l \geq |m_l|$ [16]. Likewise, the discrete energies originate from the requirement of a regular behavior of $R(0)$ and $R(\infty)$ solving Eq. (2.7).

$$E = \mathscr{E}_n = -\frac{\mu Z^2 e^4}{2\hbar^2 n^2} = -\frac{\mu}{m_e}\frac{E_{\text{Ha}}}{2}\frac{Z^2}{n^2}, \quad n = l+1,\, l+2,\, l+3, \ldots \tag{2.10}$$

The integer numbers m_l (magnetic quantum number), l (azimuthal q.n.), and n (principal q.n.) parameterize the eigenvalues and the corresponding eigenfunctions:

$$\Phi_{m_l}(\varphi) = \frac{1}{\sqrt{2\pi}}\, e^{i m_l \varphi} \tag{2.11}$$

$$\Theta_{l m_l}(\theta) = (-1)^{\frac{|m_l|+m_l}{2}} \sqrt{\frac{2l+1}{2}\frac{(l-|m_l|)!}{(l+|m_l|)!}}\, P_l^{|m_l|}(\cos\theta) \tag{2.12}$$

$$R_{n\,l}(r) = k^{3/2}\sqrt{\frac{(n-l-1)!}{2n[(n+l)!]^3}}\,(kr)^l\, L_{n+l}^{2l+1}(kr)\, e^{-kr/2}, \tag{2.13}$$

where k is a shorthand for $2Z/(an)$, a is a mass-corrected Bohr atomic length unit $a = a_0\, m_e/\mu = \hbar^2/(\mu e^2)$. For nonnegative integers l and m, the associated Legendre functions $P_l^m(x)$ are defined by

$$P_l^m(x) = (1-x^2)^{m/2}\frac{d^m}{dx^m}P_l(x), \qquad P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2-1)^l. \tag{2.14}$$

The associated Laguerre polynomials $L_p^q(\rho)$ are polynomials of degree $p-q$, defined by

$$L_p^q(\rho) = \frac{d^q}{d\rho^q}L_p(\rho), \qquad L_p(\rho) = e^\rho \frac{d^p}{d\rho^p}(\rho^p e^{-\rho}). \tag{2.15}$$

Since each individual wavefunction of Eqs. (2.11)–(2.13) is properly normalized by its own square-root factor, so is the total atomic wavefunction

$$\psi_{nlm_l}(r,\theta,\varphi) = R_{n\,l}(r)\,\Theta_{lm_l}(\theta)\,\Phi_{m_l}(\varphi) \tag{2.16}$$

representing the atomic state $|n, l, m_l\rangle$. Explicitly, the *orthonormality* relation reads:

$$\begin{aligned}\langle n, l, m_l | n', l', m_l'\rangle &= \int r^2 dr\ \sin\theta d\theta\, d\varphi\ \psi^*_{nlm_l}(r,\theta,\varphi)\,\psi_{n'l'm_l'}(r,\theta,\varphi)\\ &= \delta_{nn'}\,\delta_{ll'}\,\delta_{m_l m_l'}.\end{aligned} \tag{2.17}$$

In addition to all these *bound states*, a continuum of *unbound states* of arbitrary positive energy represents the *ionic* states, where the electron moves far away from the nucleus.

2.1.1 The Energy Spectrum

The energy eigenvalues (2.10) of the nonrelativistic one-electron atom depend on the principal quantum number n only, and exhibit the characteristic structure sketched in Fig. 2.1. In particular, the lowest-energy state is $|n, l, m_l\rangle = |1, 0, 0\rangle$. For hydrogen ($Z = 1$), its energy $\mathscr{E}_1 = -0.5\, E_{\mathrm{Ha}}\, \mu/m_e = -13.5983$ eV. Due to the reduced-mass correction $\mu/m_e = 0.999456$, this energy is slightly less negative than $-0.5\, E_{\mathrm{Ha}} = -13.6057$ eV.

Above this *ground state* we find a sequence of energy levels. The lowest excited ($n = 2$) level is 4-fold degenerate: it consists of $|2, 0, 0\rangle$, $|2, 1, -1\rangle$, $|2, 1, 0\rangle$, and $|2, 1, 1\rangle$. Its energy is $\mathscr{E}_2 = -\frac{1}{2}(Z/2)^2 E_{\mathrm{Ha}}\, \frac{\mu}{m_e} = -\frac{1}{8}\, Z^2 E_{\mathrm{Ha}}\, \frac{\mu}{m_e}$. In hydrogen this level is $\mathscr{E}_2 - \mathscr{E}_1 = \frac{3}{8}\, E_{\mathrm{Ha}}\, \frac{\mu}{m_e} = 10.1987$ eV above the ground state. Successive $\mathscr{E}_3$, $\mathscr{E}_4$… levels exhibit an increasing *degeneracy*, because of the multiple values $l = 0, \ldots, n-1$ compatible with larger n, and the values $m_l = 0, \pm 1, \ldots, \pm l$ compatible with larger l. The m_l-degeneracy ($2l + 1$ states) is a general feature of central potentials, representing the possibility for the orbital angular momentum to point in any direction in 3D space without affecting the energy of the atom. In contrast, the extra l-degeneracy (n values of l, for a total of n^2 states) is peculiar of the Coulombic potential energy $U(r) \propto -r^{-1}$: none of this "accidental" degeneracy occurs for a different radial dependence of $U(r)$, e.g. as encountered below in Sect. 2.2.4 for many-electron atoms, and in Sect. 3.3.1 for diatomic molecules.

Transitions between any two energy levels are observed, Fig. 2.2. Historically, the close agreement of the H-atom Schrödinger spectrum of Eq. (2.10) with accurate spectroscopic observations marked an early triumph of QM. The transitions group naturally in *series* of transitions $n_i \to n_f$, characterized by the same *lower*-energy level (usually the final level n_f in emission experiments, see Fig. 2.2). Each series of transitions of hydrogen is observed in a characteristic spectral region and is named after the scientist who carried out its earliest investigations: the transitions

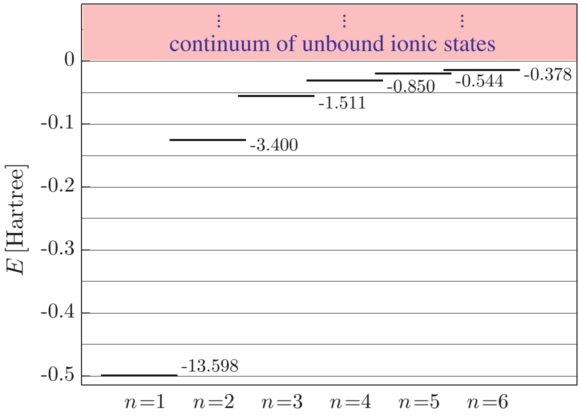


Fig. 2.1 The 6 lowest-energy levels $\mathscr{E}_n$ of atomic hydrogen, according to the nonrelativistic Schrödinger theory, Eq. (2.10). Energies in eV are indicated next to each line. The energy zero marks the onset of the continuum of unbound states

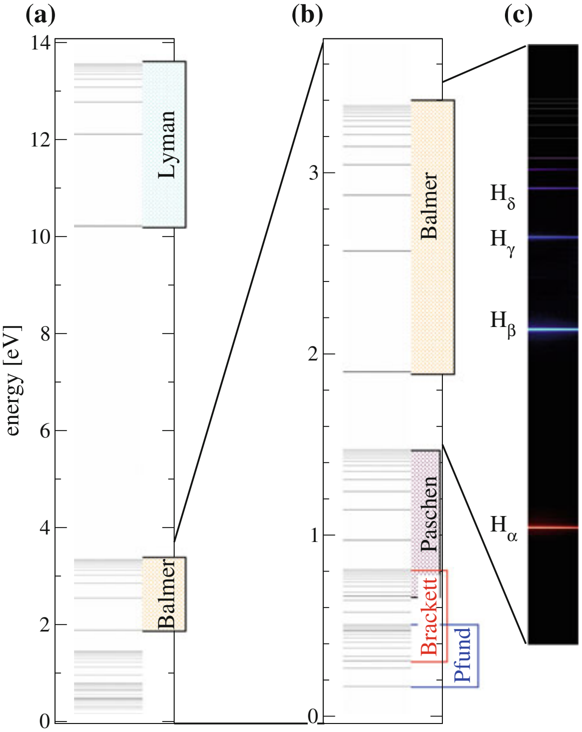


Fig. 2.2 The observed emission line spectrum of atomic hydrogen, with **a** a broad full IR-UV range spectrum, **b** a closeup of the Balmer plus the overlapping infrared series, and **c** a further closeup of the Balmer series, with its 4 visible lines, plus a few low-energy UV lines

whose lower level is $n = 1$ form the Lyman series (10.2–13.6 eV, UV); those whose lower level is $n = 2$ constitute the Balmer series (1.89–3.40 eV, visible-UV); those whose lower level is $n = 3$ form the Paschen series (0.66–1.51 eV, IR); those whose lower level is $n = 4$ form the Brackett series (0.31–0.85 eV, IR). Note in particular, that the Lyman and Balmer series do not overlap with any other series, since the energy distance between their lower level ($n = 1$ or 2) and the next one ($n + 1$) exceeds the whole range of bound-state energies from level $n + 1$ to the ionization threshold.

The weak dependence of the spectrum on the nuclear mass M [through the ratio $\mu/m_e = (1 + m_e/M)^{-1}$] produces a duplication (relative energy separation $\simeq$0.03 %—see Fig. 2.3) of the spectral lines of a mixture of different *isotopes* such as regular hydrogen ^{1}H and twice as heavy *deuterium* ^{2}H (or D). Finally, note that the Z^2 dependence of the eigenvalues Eq. (2.10) makes one half of the lines of one half of the series of the He$^+$ ion (one third of those of Li^{2+}, …) almost coincident (except for the reduced-mass correction, and relativistic effects) with the lines of hydrogen, as illustrated in Fig. 2.4.

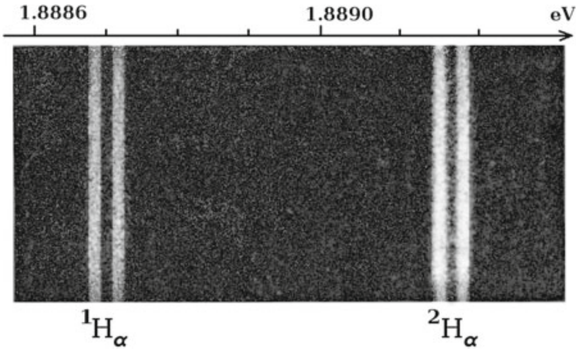


Fig. 2.3 A high-resolution line spectrum of the Balmer H_α emission of an isotope mixture. ^{1}H emits the longer-wavelength structure near 1.8887 eV; deuterium ^{2}H emits the shorter-wavelength lines near 1.8892 eV. The finer doublet structure is due to relativistic effects analyzed in Sect. 2.1.7

2.1.2 The Angular Wavefunction

The angular solutions (2.11) and (2.12) combine to form the *spherical harmonics* $Y_{lm_l}(\theta, \varphi) = \Theta_{lm_l}(\theta)\, \Phi_{m_l}(\varphi)$, which are the normalized eigenfunctions[2] of the angular motion of a freely rotating quantum-mechanical particle. Y_{lm_l} contains complete information about an important observable: the *orbital angular momentum*. In detail, $\hbar^2\times$ [the angular part of $-\nabla^2$ occurring at the left side of Eq. (2.6)] represents the squared orbital angular momentum $|\mathbf{L}|^2$ of the rotating two-body system. $\hbar^2\lambda = \hbar^2 l(l+1)$ are the eigenvalues of $|\mathbf{L}|^2$. Likewise, $\hbar m_l$ are the eigenvalues of the angular-momentum component L_z represented by $-i\hbar\frac{\partial}{\partial\varphi}$, thus clarifying the physical meaning of Eq. (2.5). Only statistical information about the L_x and L_y components is available, since these other components do not commute with L_z. Note however that the choice of the $\hat{\mathbf{z}}$ direction in 3D space (related to the choice of the polar coordinate system) is arbitrary: due to spherical symmetry, any alternative choice would lead to the same observable results.

The spherical harmonics carry complete information about the angular distribution of $\mathbf{r}$. In a state $|l, m_l\rangle$ with squared angular momentum fixed by l and z-projection by m_l, the probability that the vector $\mathbf{r}$ joining the nucleus to the electron is directed as $(\sin\theta\,\cos\varphi,\ \sin\theta\,\sin\varphi,\ \cos\theta)$ equals $|\langle\theta, \phi|l, m_l\rangle|^2 \sin\theta\, d\theta\, d\varphi \equiv |Y_{lm_l}(\theta, \varphi)|^2 \sin\theta\, d\theta\, d\varphi$. Equation (2.11) indicates that the φ dependence of $|Y_{lm_l}|^2$ is always trivial: $|Y_{lm_l}|^2$ are constant functions on all circles at fixed θ. Figure 2.5 illustrates the θ dependence as polar plots of $|Y_{lm_l}|^2$ for several values of l and m_l. Several textbooks and web sites report alternative visualizations of the Y_{lm_l} functions. It is apparent that $l - |m_l|$ counts the number of zeros (*nodes*) of $|Y_{lm_l}|^2$ as the polar angle θ spans the 0 to π range. A large number of nodes indicates a large angular-momentum component perpendicular to $\hat{\mathbf{z}}$.

Observations: (i) The visible (Fig. 2.5) increase of $|Y_{l0}(0, \varphi)|^2$ with l does not contradict normalization (2.17), because of the $\sin\theta$ integration factor. (ii) If the r^l factor taken from the radial wavefunction (2.13) is grouped together with $Y_{lm_l}(\theta, \varphi)$,

---

[2] We stick to the standard convention for the phases of $Y_{lm_l}(\theta, \varphi)$.

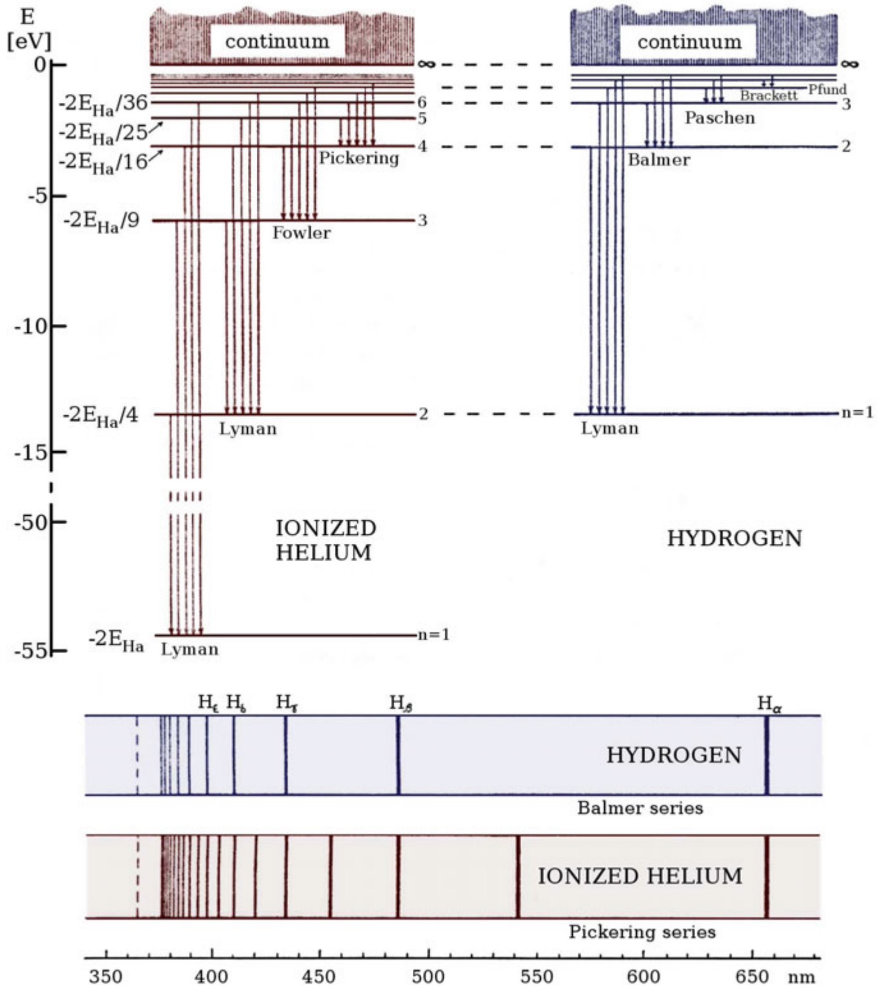


Fig. 2.4 Comparison of the emission spectra of one-electron atoms/ions with different nuclear charge: H ($Z = 1$) and He$^+$ ($Z = 2$). The odd-n series of ionized helium find no correspondence in the spectrum of hydrogen; in contrast, one half of the lines of the even-n series almost coincide with the lines of hydrogen

one can express $r^l\, Y_{lm_l}(\theta, \varphi)$ in Cartesian components $(r_x,\, r_y,\, r_z)$, obtaining a homogeneous polynomial of degree l. For example,

$$r\, Y_{10}(\theta, \varphi) = \sqrt{\frac{3}{4\pi}}\, r_z, \quad r\, Y_{1\,\pm 1}(\theta, \varphi) = \sqrt{\frac{3}{8\pi}}\, (\mp r_x - i\, r_y). \tag{2.18}$$

(iii) Observation (ii) implies that the parity of $Y_{lm_l}(\theta, \varphi)$ (i.e. its character for $\mathbf{r} \to -\mathbf{r}$) is the same as that of l, i.e. $(-1)^l$. (iv) $m_l = 0$ spherical harmonics are real

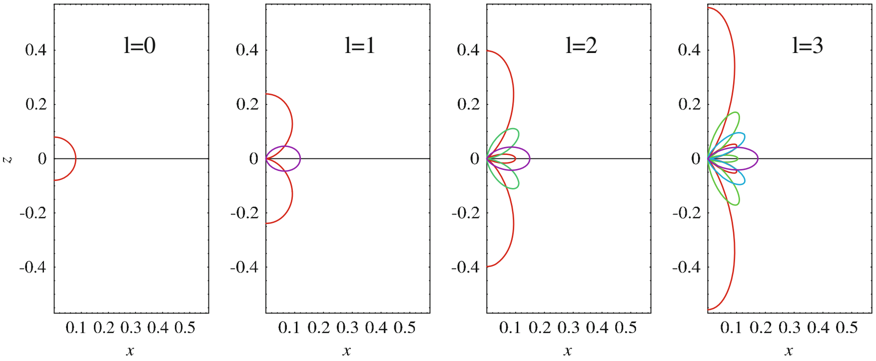


Fig. 2.5 Polar plots of the lowest-l spherical harmonics: the radial distance from the origin equals $\left|Y_{lm_l}(\theta,\varphi)\right|^2$. Here the x–z plane ($\varphi = 0$) is shown, but φ may be taken of any value. θ measures the angle away from the $\hat{\mathbf{z}}$ axis and varies from 0 (*upward*) to π (*downward*). Colors encode increasing value of $|m_l|$, from its minimum $m_l = 0$ (*red*) to its maximum $m_l = l$ (*violet*)

functions. (v) Occasionally $m_l \neq 0$ spherical harmonics are combined to construct real wavefunctions, e.g.

$$\psi_{\mathrm{p}_x} = \frac{Y_{1\,-1} - Y_{1\,1}}{\sqrt{2}}, \quad \psi_{\mathrm{p}_y} = i\,\frac{Y_{1\,1} + Y_{1\,-1}}{\sqrt{2}}. \tag{2.19}$$

(vi) It is easy and useful to retain the expression for the simplest spherical harmonic function (a polynomial of degree 0, i.e. a constant): $Y_{00}(\theta,\varphi) = (4\pi)^{-1/2}$.

Important notation: Spectroscopists have adopted a letter code for the value of the orbital angular momentum. This standard notation is: s, p, d, f, g, h,..., for $l = 0$, 1, 2, 3, 4, 5... respectively.

2.1.3 The Radial Wavefunction

The radial wavefunction $R_{n\,l}(r)$, Eq. (2.13), is structured as the product of (i) a normalization factor, (ii) a power r^l (mentioned above in relation to Y_{lm_l}), (iii) a polynomial of degree $n - l - 1$ in r (with a nonzero r^0 term), and (iv) the exponential of $-Zr/(2an)$. The power term is responsible for the $R_{n\,l}(r) \propto r^l$ behavior at small r. The Laguerre polynomial $L_{n+l}^{2l+1}(\rho)$ vanishes at as many different $\rho > 0$ points as its degree $(n - l - 1)$: each of these zeroes produces a *radial node*, i.e. a spherical shell of radius r where $R_{n\,l}(r)$, and thus the overall wavefunction, vanishes and changes sign. The exponential decay dominates at large r, where $R_{n\,l}(r) \sim r^{n-1} \exp(-Zr/na)$. Figure 2.6 illustrates these features for the square moduli of the lowest-n radial eigenfunctions.

The *radial probability distribution* $P(r) = r^2|R(r)|^2$ drawn in Fig. 2.7 and the $|R(r)|^2$ of Fig. 2.6 contain different information about the electron state. $P(r)\,dr$ yields the probability that the nucleus-electron distance is within dr of r, regardless

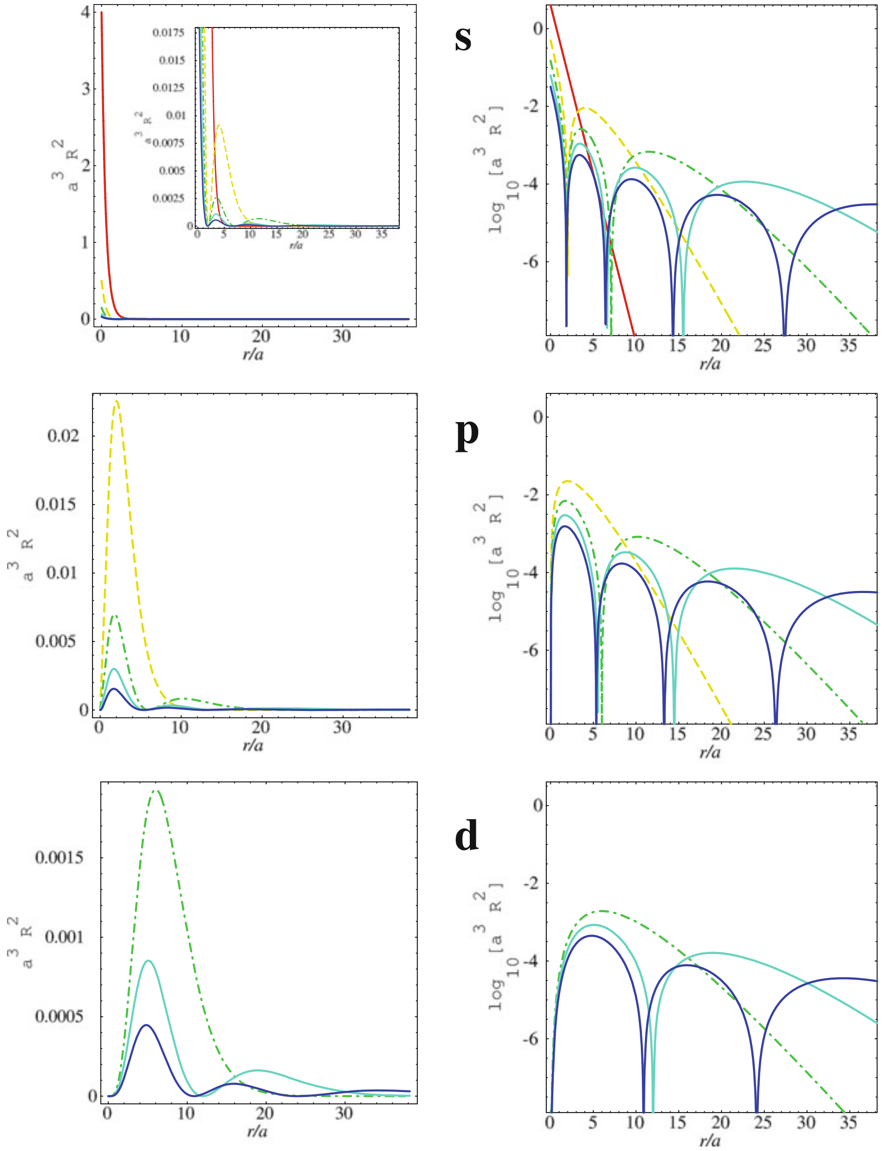


Fig. 2.6 $Z = 1$ hydrogenic s p and d squared radial eigenfunctions $|R_{n\,l}(r)|^2$. *Solid*, $n = 1$; *dashed*, $n = 2$; $\cdot -$, $n = 3$; $\cdot\cdot -$, $n = 4$; $\cdot\cdot\cdot -$, $n = 5$. Due to squaring, the $(n - l - 1)$ radial nodes, where $R_{nl}(r)$ vanishes and changes sign, appear as tangencies to the horizontal axis of the linear plots (*left*), and as downward kinks in the plots in $\log_{10}$ scale (*right*)

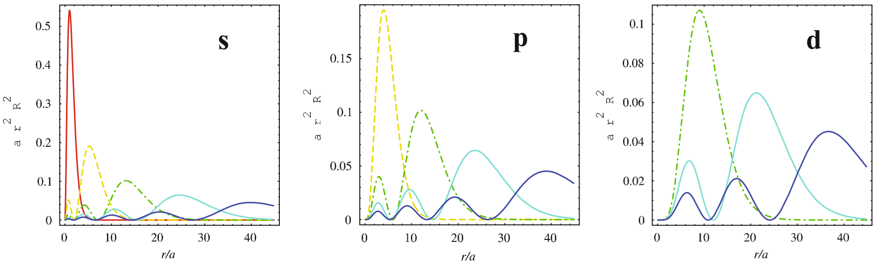


Fig. 2.7 Hydrogen ($Z = 1$) s p and d radial probability distribution $P(r) = |r\, R_{nl}(r)|^2$. *Solid*, $n = 1$; *dashed*, $n = 2$; $\cdot -$, $n = 3$; $\cdot\cdot -$, $n = 4$; $\cdot\cdot\cdot -$, $n = 5$

of the direction where $\mathbf{r}$ points. The r^2 weighting factor is precisely the spherical-coordinates Jacobian, proportional to the surface of the sphere of radius r, or rather the volume of the spherical shell "between r and $r + dr$". In Fig. 2.7, note that for increasing n, $P(r)$ peaks at larger and larger distance from the origin.

The probability that the electron is found at a specific position $\mathbf{r}$ relative to the nucleus is not $P(r)$, but is given by

$$P_{3D}(\mathbf{r})\, d^3\mathbf{r} = |\psi(\mathbf{r})|^2\, d^3\mathbf{r} = |\psi_{nlm_l}(r, \theta, \varphi)|^2\, d^3\mathbf{r} = |R_{n\,l}(r)|^2\, |Y_{lm_l}(\theta, \varphi)|^2\, d^3\mathbf{r}, \tag{2.20}$$

where the polar coordinates are those representing that point $\mathbf{r}$. Equation (2.20) indicates that $|\psi(r, \theta, \varphi)|^2$ gives the actual 3D probability distribution in space. This means, e.g., that the probability density profile along a line through the nucleus (specified by fixing θ and φ) is simply $|R(r)|^2$, multiplied by the constant $|Y_{lm_l}(\theta, \varphi)|^2$. Note that all s eigenfunctions have nonzero $R_{n\,0}(0) = 2[Z/(an)]^{3/2}$. Moreover, $r = 0$ is a cusp-type absolute *maximum* of $|R_{n\,0}(r)|^2$, see Fig. 2.6. It is no surprise that the most likely point in space for s electrons coincides with the nucleus, $\mathbf{r} = \mathbf{0}$, the spot with the most attractive potential energy $U(\mathbf{r})$. This fact is hidden by the vanishing of $P(r \to 0)$ which is entirely due to the r^2 weight in $P(r)$. For $l > 0$, even the probability density $|R_{n\,l}(r)|^2$ vanishes at the origin, where the centrifugal repulsion $\propto \lambda/r^2 = l(l+1)/r^2$ "effective potential" term in Eq. (2.7) diverges. The vanishing of $|R_{n\,l}(r)|^2$ reflects the impossibility of a point particle carrying nonzero angular momentum to reach the origin of a central potential. The wave-mechanical reason is the following: for $l > 0$, $\mathbf{r} = \mathbf{0}$ is a common point of one or several nodes of the angular wavefunction, therefore if ψ did not vanish there, it would be multiple-valued.

For increasing nuclear charge Z, $|R_{n\,l}(r)|^2$ and thus $P(r)$ move in closer and closer to the origin[3]: the mean electron–nucleus separation decreases as $\propto Z^{-1}$, and this fact combined with $V_{ne} \propto Z/r$ explains the $\propto Z^2$ dependence of the eigenenergies (2.10). The simplest radial wavefunction, that of the ground state, exemplifies well this Z dependence:

---

[3] The scaling laws are $R_{n\,l\,Z}(r) = Z^{3/2} R_{n\,l\,1}(rZ)$ and $P_Z(r) = Z\, P_1(rZ)$.

$$R_{10}(r) = \sqrt{\frac{k^3}{2}}\, e^{-kr/2} = 2\left(\frac{Z}{a}\right)^{3/2} \exp\left(-\frac{Zr}{a}\right). \qquad (2.21)$$

Accordingly, the ground state of He^+ has the same overall shape but half the size of that of H.

Notation: the hydrogenic kets/eigenfunctions $|n, l, m_l\rangle$ of Eq. (2.16) are often shorthanded as $n[l]$, where n is the principal quantum number and $[l]$ is the relevant letter s, p, d, … for that value of l. For example, 4p refers to any of $\psi_{41\,-1}$, $\psi_{41\,0}$, $\psi_{41\,1}$. This notation is incomplete and ambiguous: (i) information about m_l is lacking, and (ii) the same 4p symbol implies different radial dependence $R_{41}(r)$ for a nucleus of different charge and mass.

2.1.4 Orbital Angular Momentum and Magnetic Dipole Moment

The angular momentum of an orbiting charged particle such as an electron is associated to a magnetic dipole moment. This is illustrated (Fig. 2.8) for a classical point particle of mass m and charge q rotating along a circular orbit of radius r at speed v. Its angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p} = mrv\hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is the unit vector perpendicular to its trajectory. As the rotation period is $\tau = 2\pi r/v$, the current along the loop $I = q/\tau = q\,v/(2\pi r)$. The magnetic moment of a ring current equals the product of the current times the loop area:

$$\boldsymbol{\mu} = I\,\pi r^2 \hat{\mathbf{n}} = \frac{qv}{2\pi r}\pi r^2 \hat{\mathbf{n}} = \frac{q}{2}\, v\, r\, \hat{\mathbf{n}} = \frac{q}{2m}\mathbf{L}. \qquad (2.22)$$

One can show that this equality holds for arbitrary shape of the periodic orbit.

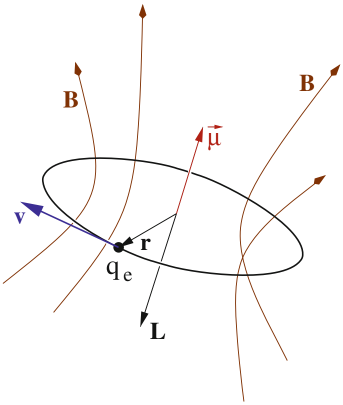


Fig. 2.8 The relation between the mechanical angular momentum **L** and the magnetic moment $\boldsymbol{\mu}$ generated by an electron of charge $-q_e$ orbiting circularly. The *curved lines* represent the magnetic induction field **B** produced by the circulating current

Relation (2.22) holds also in QM, as an operatorial relation. For an electron of charge $q = -q_e$, where the angular momentum is quantized in units of $\hbar$, it is convenient to write Eq. (2.22) as

$$\boldsymbol{\mu} = -\frac{q_e}{2m_e}\,\mathbf{L} = -\frac{\hbar q_e}{2m_e}\,\frac{\mathbf{L}}{\hbar} = -g_l\,\mu_\mathrm{B}\,\frac{\mathbf{L}}{\hbar}, \tag{2.23}$$

where the *Bohr magneton* $\mu_\mathrm{B} = \hbar q_e/(2m_e) = 9.27401 \times 10^{-24}\,\mathrm{J\,T^{-1}}$ (alias $\mathrm{A\,m^2}$) is the natural scale of atomic magnetic moments. $g_l = 1$ is the orbital g-factor, introduced for uniformity of notation with those situations, discussed below, with nontrivial proportionality factors $g \neq 1$ between $\mu_\mathrm{B}\,\mathbf{L}/\hbar$ and $\boldsymbol{\mu}$. g-factors arise because, while angular momenta are universal (i.e. simple multiples of $\hbar$), magnetic moments are usually non-universal multiples of μ_B.

The atomic angular momenta can be detected by letting the associated magnetic moments interact with a magnetic field. If the field $\mathbf{B}$ is uniform, it induces a precession of $\boldsymbol{\mu}$ around the direction of $\mathbf{B}$ with a frequency (the Larmor frequency) $\omega = q_e B/(2m_e)$, routinely detected in microwave resonance experiments. If the field is nonuniform instead, a net force acts on the atom, as we discuss in the next section.

2.1.5 The Stern-Gerlach Experiment

The interaction energy of a magnetic moment with a magnetic field is

$$H_\mathrm{magn} = -\boldsymbol{\mu} \cdot \mathbf{B}. \tag{2.24}$$

Unless some external mechanism alters the angle between $\boldsymbol{\mu}$ and $\mathbf{B}$, this energy is conserved in time. A force arises on the magnetic dipole when the field $\mathbf{B}$ is nonconstant in space:

$$\mathbf{F} = -\boldsymbol{\nabla}(-\boldsymbol{\mu} \cdot \mathbf{B}) = \boldsymbol{\nabla}(\boldsymbol{\mu} \cdot \mathbf{B}). \tag{2.25}$$

In particular, in a magnetic field with a dominant B_z component, the z force component F_z is proportional to the derivative of B_z along the same direction:

$$F_z \simeq \mu_z\, \nabla_z B_z = \mu_z\, \frac{\partial B_z}{\partial z}. \tag{2.26}$$

The “microscopical” origin of this force is pictured in Fig. 2.9. The observation that a nonuniform magnetic field produces a force proportional to a magnetic-moment component is at the basis of the Stern-Gerlach experiment.

As illustrated in Fig. 2.10, a collimated beam of neutral atoms at thermal speeds is emitted from an oven into a vacuum chamber where it traverses a region of

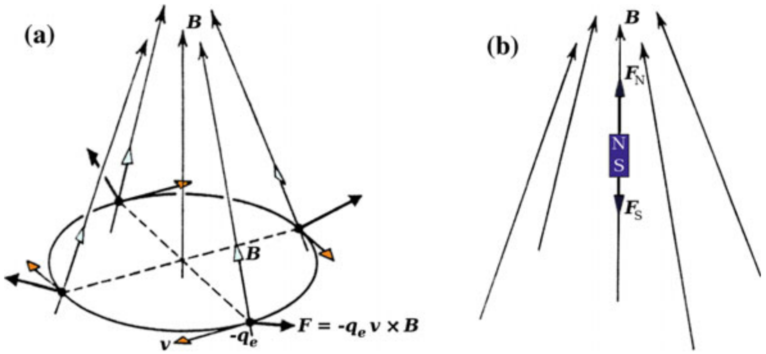


Fig. 2.9 The origin of the force that a nonuniform magnetic field produces on a magnetic moment. **a** If the magnetic dipole is seen as a circulating current, the net force originates from a force component consistently pointing in the direction of increasing **B**. **b** If the dipole is viewed as a pair magnetic monopoles, a net force arises from the unbalance between the forces on the individual monopoles

inhomogeneous magnetic field, where F_z deflects individual atoms proportionally to their magnetic moment z component, and is finally collected by a suitable detector. Basically, the Stern-Gerlach apparatus is an instrument for measuring the component of atomic magnetic moments in the field-gradient direction. This makes it one of the key tools in QM, as detailed in the initial sections of Ref. [17]. The original (1922) experiment was carried out using Ag atoms, but a similar pattern of deflections is observed using atomic H.

The main outcome of the Stern-Gerlach experiment is that the z component of $\boldsymbol{\mu}$ is not distributed continuously as one would expect for a classical vector pointing randomly in space, but rather peaked at discrete values. Figure 2.10c shows the clustering of the deflected atoms in two lumps.

According to QM, the $\hat{\mathbf{z}}$-component L_z of angular momentum (and thus μ_z of magnetic moment) should indeed exhibit discrete eigenvalues. However, (i) the number of eigenvalues of L_z should be odd [$2l + 1$, with integer l—see Eq. (2.9)] and (ii) the ground state of hydrogen has $l = 0$, thus H should show no magnetic moment at all, and one undeflected lump should be observed, rather than splitting into two. This is a first hint that some extra degree of freedom must play a role in the one-electron atom.

2.1.6 *Electron Spin*

The outcome of the Stern-Gerlach experiment, the multiplet fine structure of the spectral lines (the fine doublets of Fig. 2.3), and the Zeeman splitting of the spectral lines (see Sect. 2.1.10) are three pieces of evidence pointing at the existence of an extra degree of freedom of the electron, beside its position in space. W. Pauli introduced

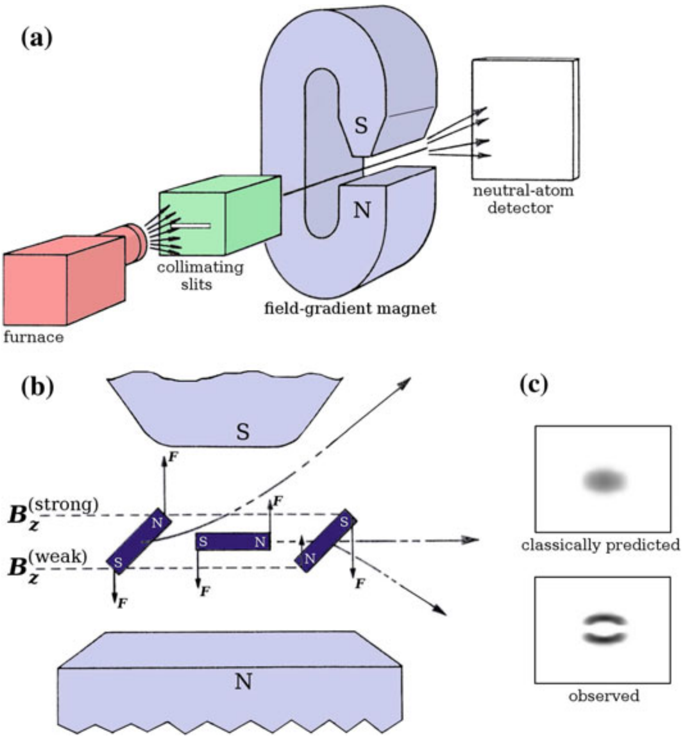


Fig. 2.10 **a** In the Stern-Gerlach experiment, a collimated beam of atoms emitted from an oven transverses a region of inhomogeneous magnetic field created by a magnet with asymmetric core expansions: the atoms are finally detected at a collector plate. **b** In an inhomogeneous magnetic field, a magnet experiences a net force which depends on its orientation. **c** The deflection pattern recorded on the detecting plate in a Stern-Gerlach measurement of the z component of the magnetic dipole moment of Ag atoms (the outcome would be the same for H atoms). Contrary to the classical prediction of an even distribution of randomly oriented magnetic moments, two discrete components are observed, due to the quantization of an angular-momentum component

a nonclassical internal degree of freedom, later named *spin*, with properties similar to orbital angular momentum. Even though this picture is imprecise, the electron spin may be viewed as the intrinsic angular momentum of the rotation of the electron around itself. When spin is measured along a given direction, say $\hat{\mathbf{z}}$, one detects eigenvalues of S_z, $\hbar m_s$, where the quantum number m_s takes $(2s+1)$ values $m_s = -s, \ldots, s$, like the orbital L_z takes values $m_l = -l, \ldots, l$. Since a Stern-Gerlach deflector splits an H beam into *two* lumps, $2 = 2s + 1$ components are postulated, requiring that the intrinsic angular momentum of the electron must be $s = 1/2$. This in turn is associated to a squared spin angular-momentum operator $|\mathbf{S}|^2$ whose eigenvalue is $1/2(1/2 + 1)\hbar^2 = 3/4\hbar^2$.

A complete wavefunction, necessary to specify all degrees of freedom of the electron, is slightly more complicated than $R(r)\, Y_{lm_l}(\theta, \varphi)$: an extra spin dependence must be inserted. Assuming, as apparent from the nonrelativistic Hamiltonian (1.1), that spin and orbital motions do not interact, the full eigenfunction of a one-electron atom with spin pointing up (↑, i.e. $m_s = 1/2$) or down (↓, i.e. $m_s = -1/2$) along a fixed orientation is written

$$\psi_{n\,l\,m_l m_s}(r, \theta, \varphi, \sigma) = R_{n\,l}(r)\, Y_{l\,m_l}(\theta, \varphi)\, \chi_{m_s}(\sigma). \tag{2.27}$$

Here σ is the variable for the spin degree of freedom, which spans values $\pm 1/2$, i.e. checking if the electron spin points up or down in the $\hat{\mathbf{z}}$ direction. The quantum number $m_s = \pm 1/2$ indicates which way the spin of this specific state is actually pointing relative to the reference direction. These basis spin functions are therefore simply $\chi_{m_s}(\sigma) = \langle \sigma | m_s \rangle = \delta_{m_s\,\sigma}$.

Less trivial spin wavefunctions arise when the spin points in some direction other than $\hat{\mathbf{z}}$ (non-S_z eigenstates). A Stern-Gerlach apparatus can be employed to purify a *spin polarized* beam of atoms with spins pointing in some oblique direction. This beam can then be analyzed by a second apparatus to measure the spin component σ along the reference $\hat{\mathbf{z}}$ direction. For the oblique-spin state, the (now nontrivial) spin wavefunction $\chi(\sigma)$ bears the standard significance of a wavefunction in QM: $|\chi(\uparrow)|^2$ is the probability that, when S_z is measured, $+1/2\hbar$ is found, while $|\chi(\downarrow)|^2$ is the probability to obtain $-1/2\hbar$. When a $\hat{\mathbf{z}}$-Stern-Gerlach measurement is carried out, the spin z component is found pointing either up or down, therefore the total probability $\sum_\sigma |\chi(\sigma)|^2 = |\chi(\uparrow)|^2 + |\chi(\downarrow)|^2 = 1$.

A remarkable novelty regarding spin is that the separation of the ↑ and ↓ sub-beams in a Stern-Gerlach apparatus is compatible with a g-factor for spin $g_s \simeq 2$, quite distinct from the orbital $g_l = 1$. The precise value $g_s = 2.00232$ generating the electron intrinsic magnetic moment $g_s \mu_B$ is measured extremely accurately by electron spin resonance (ESR) experiments, where the electron spin interacts with a uniform magnetic field which splits the ↑ and ↓ states.

At the present basic level of understanding, electron spin is just an extra quantum number which, in the absence of magnetic fields only provides an extra degeneracy to all states of the one-electron atom: the total degeneracy of the nth level is $2n^2$, rather than n^2. Spin will however affect the atomic energy levels when the magnetic effects of relativistic origin are considered, in Sect. 2.1.7.

2.1.7 Fine Structure

The smallness of the observed fine splittings (a fraction of meV, see Fig. 2.3) in the spectrum of H hints at some small correction, such as due to relativistic effects, neglected in the original Hamiltonian (1.1). We come to investigate these corrections in detail.

2.1.7.1 Spin-Orbit Coupling

Consider first the action of the magnetic field experienced by the electron spin due to its own orbital motion. This is a subtle relativistic effect, due to the Lorentz transformation of the nuclear electric field into the frame of reference of the electron. Call **v** the electron velocity in the nuclear rest frame. In the electron frame of reference, the nucleus is seen to move with velocity $-\mathbf{v}$, and thus carries a "current element" $-Zq_e\mathbf{v}$. According to the Biot-Savart law of electromagnetism, at the point where the electron sits (reached from the nucleus by vector **r**), this moving charge generates a magnetic field

$$\mathbf{B}(\mathbf{r}) = -\frac{1}{4\pi\varepsilon_0 c^2}\frac{\mathbf{r}\times(-Zq_e\mathbf{v})}{|\mathbf{r}|^3} = \frac{\mathbf{E}(\mathbf{r})\times\mathbf{v}}{c^2}. \tag{2.28}$$

Equation (2.28) identifies this magnetic field as a relativistic effect [of order $(v/c)^2$], and expresses it in terms of the electric field generated by the nucleus at the electron location

$$\mathbf{E}(\mathbf{r}) = \frac{Zq_e}{4\pi\varepsilon_0}\frac{\mathbf{r}}{|\mathbf{r}|^3}. \tag{2.29}$$

In Eq. (2.28) we recognize the orbital angular-momentum operator:

$$\mathbf{B}(\mathbf{r}) = \frac{Zq_e}{4\pi\varepsilon_0 c^2}\frac{\mathbf{r}\times\mathbf{v}}{|\mathbf{r}|^3} = \frac{Zq_e}{4\pi\varepsilon_0 c^2 m_e}\frac{\mathbf{L}}{|\mathbf{r}|^3}. \tag{2.30}$$

This magnetic field $\mathbf{B}(\mathbf{r})$ acts on the electron at each point in space. By analogy with Eq. (2.24), the interaction energy of the electron-spin magnetic moment $\boldsymbol{\mu}_s$ with **B** should equal $-\boldsymbol{\mu}_s\cdot\mathbf{B}(\mathbf{r})$. However, this energy must actually be reduced by a factor 1/2 (first recognized by L.H. Thomas) due to the electron frame of reference being accelerated [1]. The correct magnetic interaction energy operator is therefore:

$$H_{\text{s-o}} = -\frac{1}{2}\,\boldsymbol{\mu}_s\cdot\mathbf{B}(\mathbf{r}) = \frac{1}{2}\,\frac{g_s\mu_{\text{B}}}{\hbar}\,\mathbf{S}\cdot\left(\frac{Zq_e}{4\pi\varepsilon_0 c^2 m_e}\,\frac{\mathbf{L}}{|\mathbf{r}|^3}\right) = \frac{Ze^2}{2\,m_e^2\,c^2}\,\frac{1}{r^3}\,\mathbf{S}\cdot\mathbf{L}. \tag{2.31}$$

This operator, named *spin-orbit interaction*, exhibits nonzero diagonal and off-diagonal elements $\langle n,l,m_l,m_s|H_{\text{s-o}}|n',l,m_l',m_s'\rangle$ connecting states with equal or different n, m_l, m_s. States with $n\neq n'$ have vastly different nonrelativistic energies (2.10): the tiny n-off-diagonal spin-orbit couplings perturb these energies negligibly, and can practically be neglected, see Appendix B.5.2 and B.9. For given l and considering only the n-diagonal matrix elements of $H_{\text{s-o}}$, we rewrite Eq. (2.31) as

$$H_{\text{s-o}} \simeq \frac{Ze^2}{2\,m_e^2\,c^2}\sum_{n,l}|n,l\rangle\langle n,l|r^{-3}|n,l\rangle\langle n,l|\,\mathbf{S}\cdot\mathbf{L}. \tag{2.32}$$

The radial integral can be evaluated for $l \gneq 0$ hydrogenic wavefunctions $R_{n\,l}$, obtaining

$$\langle n, l|r^{-3}|n, l\rangle = \int_0^\infty r^{-3}[R_{n\,l}(r)]^2\, r^2\, dr = \left(\frac{Z}{a}\right)^3 \frac{2}{n^3\, l(l+1)(2l+1)}. \tag{2.33}$$

The spin-orbit Hamiltonian is thus conveniently rewritten as

$$H_{\text{s-o}} = \sum_{n,l} \xi_{nl}\, |n, l\rangle\langle n, l|\, \frac{\mathbf{S}\cdot\mathbf{L}}{\hbar^2}, \tag{2.34}$$

where the projectors $|n, l\rangle\langle n, l|$ select a radial wavefunction, and the spin-orbit energy is

$$\xi_{nl} = \frac{Ze^2\hbar^2}{2\,m_e^2\,c^2}\left(\frac{Z}{a}\right)^3 \frac{2}{n^3\, l(l+1)(2l+1)} = Z^4\alpha^2 E_{\text{Ha}}\left(\frac{\mu}{m_e}\right)^3 \frac{1}{n^3\, l(l+1)(2l+1)}. \tag{2.35}$$

The last equality uses the expression for the mass-rescaled atomic length scale $a = a_0\, m_e/\mu$, the definition (1.9) of the Hartree energy, and the expression $\alpha^2 = E_{\text{Ha}}/(m_e c^2)$ for the fine-structure constant. In this form, it is apparent that the typical spin-orbit energy scale ξ_{nl}

- is positive, and therefore $H_{\text{s-o}}$ favors *antiparallel* alignment of **L** and **S**;
- is a leading $\alpha^2 \sim (v/c)^2$ relativistic correction;
- is $\alpha^2 \simeq 5.3 \times 10^{-5}$ times smaller than the typical orbital energies;
- grows as Z^4, reflecting the increase in nuclear field intensity $\propto Z$ and the reduction $\propto Z^{-1}$ of the average electron–nucleus distance so that $\langle n, l|r^{-3}|n, l\rangle \propto Z^3$;
- decreases as n^{-3}, reflecting the increase $\propto n$ of the average electron–nucleus distance, and the r^{-3} dependence of the interaction energy (2.31);
- decreases roughly as l^{-3} (but $\mathbf{S}\cdot\mathbf{L} \propto l$), due to the r^l suppression of the radial wavefunction close to the origin, the region where spin-orbit interaction (2.31) dominates.

The energy scale of ξ_{nl} amounts to $\alpha^2 E_{\text{Ha}}\,(\mu/m_e)^3 = 2.318 \times 10^{-22}\,\text{J} = 1.447\,\text{meV}$ for hydrogen. Note however that the lowest-energy state for which spin-orbit applies, 2p, has $n^3 l(l+1)(2l+1) = 48$, thus $\xi_{2\text{p}} = 0.0301\,\text{meV}$ only. For all higher levels ξ_{nl} is even smaller.

Consider now the remaining operatorial part in Eq. (2.34): $\mathbf{S}\cdot\mathbf{L}$. In the basis $|l,\, s,\, m_l,\, m_s\rangle$ where spin and orbital motion are uncoupled, the operator $\mathbf{S}\cdot\mathbf{L}$ has plenty of nonzero off-diagonal matrix elements. In computing the relativistic corrected eigenvalues and eigenstates with the methods of Appendix B.5, one needs to diagonalize $\mathbf{S}\cdot\mathbf{L}$, within each (initially degenerate) space at fixed l and s. Introducing the coupling of spin and orbital angular momenta with the methods of

Appendix B.8.1, the $|l,\ s,\ j,\ m_j\rangle$ coupled basis introduced in Eq. (B.73) shows its advantage: $\mathbf{S}\cdot\mathbf{L}$ is diagonal in this coupled basis. To prove this result, take the square of $(\mathbf{L}+\mathbf{S})=\mathbf{J}$ and invert it as follows:

$$\mathbf{S}\cdot\mathbf{L}=\frac{|\mathbf{J}|^2-|\mathbf{S}|^2-|\mathbf{L}|^2}{2}. \tag{2.36}$$

All operators at the right hand side are of course diagonal in the $|l,\ s,\ j,\ m_j\rangle$ basis: the expression for the general matrix element is therefore

$$\langle l,\ s,\ j,\ m_j|\frac{\mathbf{S}\cdot\mathbf{L}}{\hbar^2}|l,\ s,\ j',\ m_j'\rangle=\frac{j(j+1)-s(s+1)-l(l+1)}{2}\,\delta_{j\,j'}\,\delta_{m_j\,m_j'}. \tag{2.37}$$

In summary: in the coupled basis $|l,\ s,\ j,\ m_j\rangle$, the spin-orbit interaction is diagonal and its eigenvalues are given by Eq. (2.37), multiplied by the energy ξ_{nl}.

Notation: states of the coupled basis are commonly indicated as ${}^{2s+1}[l]_j$, where $[l]$ is the relevant capital letter S, P, D,... for that value of $l=0,1,2,\ldots$ Information about n is encoded elsewhere, e.g. in the $n[l]$ notation, and information about m_j is usually omitted. For example, 3d ${}^2D_{3/2}$ stands for any of the four $\left|n=3,\ l=2,\ j=\frac{3}{2},\ m_j\right\rangle$ kets.

As an example of spin-orbit split states, for one p level of a one-electron atom, the two different eigenvalues of the $\mathbf{S}\cdot\mathbf{L}/\hbar^2$ operator

$$\langle 1,\frac{1}{2},j,m_j|\,\frac{\mathbf{S}\cdot\mathbf{L}}{\hbar^2}\,|1,\frac{1}{2},j,m_j\rangle=\begin{cases}-1 & \text{for } j={}^1\!/_2\\ +{}^1\!/_2 & \text{for } j={}^3\!/_2\end{cases}.$$

Accordingly, spin-orbit splits any p level ($3\times 2=6$ orbital×spin states) of 1-electron atoms into a doublet ${}^2P_{1/2}$ plus a quartet ${}^2P_{3/2}$, separated by an energy ${}^3\!/_2\,\xi_{n1}$. For the 2p level of hydrogen this splitting amounts to 45.2 μeV.

2.1.7.2 The Relativistic Kinetic Correction

A second relativistic correction of the same order $(v/c)^2$ as spin-orbit must be included. This energy contribution accounts for the *leading correction to the kinetic energy* expression $p^2/(2\mu)$:

$$\begin{aligned}T_r&=\sqrt{\mu^2c^4+p^2c^2}-\mu c^2=\mu c^2\left(1+\frac{1}{2}\frac{p^2}{\mu^2c^2}-\frac{1}{8}\frac{p^4}{\mu^4c^4}+\cdots-1\right)\\&=\frac{p^2}{2\mu}-\frac{p^4}{8\mu^3c^2}+\cdots\end{aligned} \tag{2.38}$$

Like for $H_{\text{s-o}}$, to treat the weak perturbation $-p^4/(8\mu^3c^2)$ at first order, we just need the n-diagonal matrix elements of this operator. Although p^4 looks like a formidable differential operator, the trick $p^4=(p^2)^2=[2\mu(H_{\text{Coul}}-V_{ne})]^2$ allows us to rewrite the diagonal matrix elements of p^4 in terms of simple radial integrals of r^{-1} and r^{-2}. The final result is

$$\langle n, l | -\frac{p^4}{8\mu^3 c^2} | n, l \rangle = -\frac{Z^4\alpha^2}{n^3} E_{\mathrm{Ha}} \left(\frac{\mu}{m_e}\right)^3 \left(\frac{1}{2l+1} - \frac{3}{8n}\right), \tag{2.39}$$

where we omit either m_l, m_s or j, m_j, which are irrelevant for such radial integrals.

By combining the spin-orbit and kinetic correction

$$H_{\mathrm{rel}} = H_{\mathrm{s-o}} - \frac{p^4}{8\mu^3 c^2}, \tag{2.40}$$

we obtain the diagonal matrix elements of the total relativistic correction to order α^2:

$$\begin{aligned} \langle n, l, j | H_{\mathrm{rel}} | n, l, j \rangle &= \frac{Z^4\alpha^2}{n^3} E_{\mathrm{Ha}} \left(\frac{\mu}{m_e}\right)^3 \\ &\quad \times \left[\frac{j(j+1) - s(s+1) - l(l+1)}{2l(l+1)(2l+1)} - \frac{1}{2l+1} + \frac{3}{8n}\right] \\ &= -\frac{Z^4\alpha^2}{n^3} E_{\mathrm{Ha}} \left(\frac{\mu}{m_e}\right)^3 \left(\frac{1}{2j+1} - \frac{3}{8n}\right), \end{aligned} \tag{2.41}$$

where the last simplification is based on spin being $s = 1/2$, thus $l = j \pm 1/2$. A separate derivation shows that this Eq. (2.41) (but not all previous steps) holds for s states too.

Expression (2.41) can be combined with the nonrelativistic eigenvalues (2.10) to obtain the following expression for the energy eigenvalues, correct to order α^2:

$$\langle n, l, j | H_{\mathrm{tot}} + H_{\mathrm{rel}} | n, l, j \rangle = -\frac{E_{\mathrm{Ha}}}{2} \frac{\mu}{m_e} \frac{Z^2}{n^2} \left[1 + \left(Z\alpha \frac{\mu}{m_e}\right)^2 \frac{1}{n} \left(\frac{2}{2j+1} - \frac{3}{4n}\right)\right]. \tag{2.42}$$

This remarkable relation yields a quantitative prediction for the spectrum that can be directly compared to experiment: all n-levels should be split, for the different values of j, but not for different values of l giving the same j, e.g. $^2\mathrm{S}_{1/2}$ and $^2\mathrm{P}_{1/2}$ for the same n. This extra l-degeneracy is retained even in the solutions of Dirac's equation, which is exact to all orders in α, not just α^2 as Eq. (2.42).

2.1.7.3 The Lamb Shift

As the l-degeneracy is rather surprising, the occurrence of a splitting between levels with same j and different l was investigated closely, both theoretically and experimentally. Indeed, quantum fluctuations of the electromagnetic field and the finite nuclear size eventually lift this degeneracy, introducing tiny splittings named *Lamb*

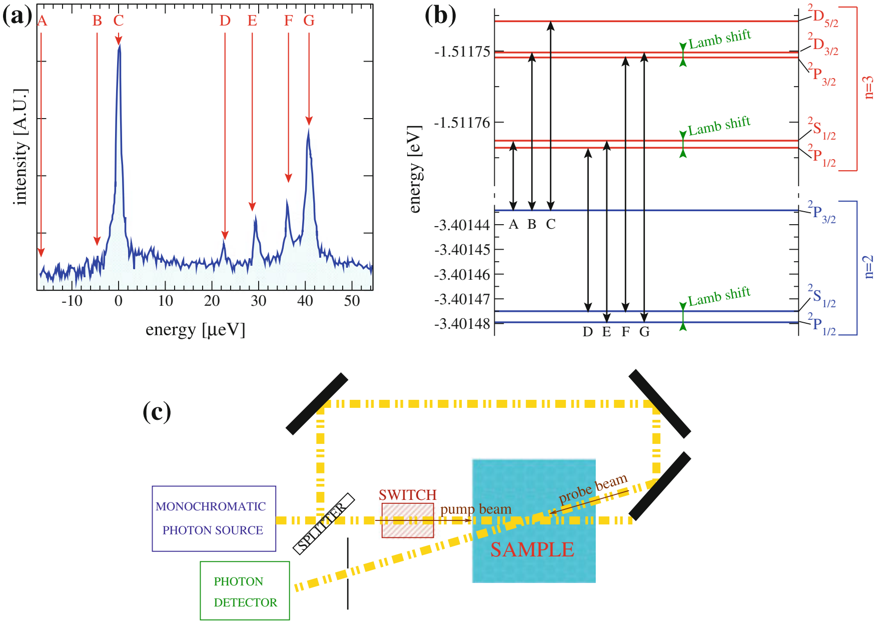


Fig. 2.11 **a** A direct spectroscopic observation of the Lamb shift in the H_α line, obtained with a double-resonance experiment [18], sketched in panel **c**. *Arrows* point to the dipole-allowed transition energies predicted according to the scheme of panel **b**, which reports the fine structure of the Balmer H_α line, including the relativistic corrections and the Lamb shift [18]. Peaks labeled *A*, *B*, and *C* involve the $n = 2\ ^2P_{3/2}$ state. Peaks labeled *D*–*G* involve the $n = 2\ j = 1/2$ states. The splitting between (*D, E*) and (*F, G*) is due to the spin-orbit coupling in the $n = 3$ multiplet. Instead, the splittings between (*D* and *E*) and between (*F* and *G*) is due to the Lamb shift, which is thus measured of the order of 5 μeV. Without the Lamb shift, the *D*–*G* peaks would collapse to 2 peaks only. The separation between the *B* and *G lines* measures the spin-orbit splitting in the 2p level: it agrees with the theoretical prediction of ≃45 μeV

shift. Figure 2.11b reports the expected spectral fine structure of the Balmer H_α line, including both the relativistic correction and the Lamb shift.

Due mainly to Doppler broadening (see Sect. 1.2), the spectral lines are not sharp enough to resolve these tiny energy differences. To circumvent Doppler broadening and acquire the high-resolution spectrum of Fig. 2.11a, the authors of Ref. [18] devised a trick based on double resonance. An intense tunable monochromatic light beam is split into a strong interruptible "pump" beam plus a second weak "probe" beam, Fig. 2.11c. When the light frequency matches a resonant transition, absorption takes place and the probe beam is attenuated, as in a regular absorption spectroscopy experiment. This absorption is strongly reduced if the pump beam happens to "saturate" the transition in the sample, as discussed quantitatively in Sect. 4.4. The spectrum of Fig. 2.11a records the probe-beam absorption *difference* between time intervals when the pump beam is on and when it is off. All atoms with a sizable translational velocity component in the beam direction are Doppler shifted in opposite directions relative to the two beams, which are almost antiparallel: these atoms do not contribute to the difference signal. As a result, a subset of atoms is selected by the matching pump/probe frequency, namely those with practically null instantaneous translational velocity component, thus negligible Doppler shift: this selection prevents Doppler broadening.

2.1.8 Nuclear Spin and Hyperfine Structure

Like electrons, several nuclei carry an intrinsic spin **I**. For example, the proton has spin $I = 1/2$. Unsurprisingly, the nuclear magnetic moment $\boldsymbol{\mu}_N$ is proportional to its angular momentum:

$$\boldsymbol{\mu}_N = g_n \mu_n \frac{\mathbf{I}}{\hbar}.$$

Like for the electron, the nuclear magneton μ_n is defined by

$$\mu_n = \frac{q_e \hbar}{2M_n} = \mu_\mathrm{B} \frac{m_e}{M_n}, \tag{2.43}$$

and the nuclear g-factor g_n is a number of order unity whose value depends on the inner nuclear structure. For example, $g_n = 5.58569$ for the proton.

The nuclear spin generates a magnetic field that is extremely weak compared to typical electronic fields, because it is suppressed by the ratio m_e/M_n. Through this field, the nuclear and electron magnetic moments interact. Due to the r^l term in Eq. (2.13), this interaction is very weak for $l > 0$ orbitals, where the electron hardly ever moves close to the nucleus: for simplicity we neglect it here. For s orbitals, the only nonzero electronic magnetic moment is associated to the electron spin **S**. Like for spin-orbit, the interaction Hamiltonian is proportional to the simplest scalar combination of the two involved vector quantities:

$$H_\mathbf{SI} = -C \boldsymbol{\mu}_N \cdot \boldsymbol{\mu}_e = C\, g_n g_s\, \mu_\mathrm{B}^2 \,\frac{m_e}{M_n}\, \frac{\mathbf{I} \cdot \mathbf{S}}{\hbar^2}.$$

The coupling factor C is the relevant radial matrix element, which equals

$$C = \frac{2}{3}\, \frac{1}{4\pi\varepsilon_0 c^2} R_{n\,0}(0)^2.$$

Using $R_{n\,0}(0) = 2[Z/(an)]^{3/2}$, the characteristic coupling energy

$$\begin{aligned} \xi_N &= C g_n g_s \mu_\mathrm{B}^2 \frac{m_e}{M_n} = \frac{2}{3}\, g_n g_s\, \frac{Z^3}{n^3 m_e c^2}\, E_\mathrm{Ha}^2 \frac{m_e}{M_n} \\ &= \frac{4}{3}\, g_n\, \frac{Z^3 \alpha^2}{n^3}\, \frac{m_e}{M_n}\, E_\mathrm{Ha} = g_n \frac{Z^3}{n^3}\, \frac{1.06\,\mu\mathrm{eV}}{M_n/\mathrm{a.m.u.}} \end{aligned} \tag{2.44}$$

where we replaced α^2 for $E_{\mathrm{Ha}}/(m_e c^2)$, and expressed the nuclear mass M_n in atomic mass units (a.m.u.). For the ground state of ^{1}H, Eq. (2.44) yields $\xi_N = 5.88\,\mu\mathrm{eV}$.

The electron and nuclear spins couple to a grand total angular momentum $\mathbf{F} = \mathbf{I} + \mathbf{S}$, with the rules described in Appendix B.8.1. In analogy with Eq. (2.36), we obtain

$$\mathbf{I} \cdot \mathbf{S} = \frac{|\mathbf{F}|^2 - |\mathbf{I}|^2 - |\mathbf{S}|^2}{2},$$

so that the expectation value of $\mathbf{I} \cdot \mathbf{S}/\hbar^2$ equals $[f(f+1) - i(i+1) - s(s+1)]/2$ in the coupled basis, the one where $|\mathbf{F}|^2$ and $|F_z|$, rather than I_z and S_z, are diagonal. As $s = 1/2$, two coupled states $f = i \pm 1/2$ occur, with an energy separation of $\xi_N\,(i + 1/2)$.

For ^{1}H, the proton has spin $i = 1/2$, so that $\mathbf{I} \cdot \mathbf{S}/\hbar^2 = -3/4$ and $1/4$ for $f = 0$ and 1, respectively. The separation between these two hyperfine-split states equals therefore $\xi_N \simeq 5.88\,\mu\mathrm{eV}$: it corresponds to a wavelength $2\pi\hbar c\,\xi_N^{-1} \simeq 21\,\mathrm{cm}$, and a frequency of $\xi_N\,(2\pi\hbar)^{-1} \simeq 1.42\,\mathrm{GHz}$. This transition, in the radio-frequency range at 1420405751.8 Hz, was discovered in 1951 in astrophysical spectra, has been adopted as a frequency standard, and is now used to map the distribution of interstellar atomic ^{1}H.

2.1.9 Electronic Transitions, Selection Rules

Not all conceivable transitions are equally easy to observe. Experimentally, certain transitions proceed at a fast rate, while others occur immensely more slowly. This fact can be explained by a quantum-mechanical analysis of the interaction of the atom with the electromagnetic field. The probability per unit time that a quantum object (e.g. an atom) decays radiatively from an initial state $|i\rangle$ to a final state $|f\rangle$ is given by

$$\gamma_{if} = \frac{1}{3\pi\varepsilon_0\hbar^4 c^3}\,\mathscr{E}_{if}^3\,|\langle f|\mathbf{d}|i\rangle|^2, \tag{2.45}$$

where $\mathscr{E}_{if} = \hbar\omega_{if} = E_i - E_f$, and $\mathbf{d}$ is the operator describing coupling to the radiation field. In the approximation that the radiating object is much smaller than the radiation wavelength (see Fig. 1.2), this operator is the electric-dipole operator $\mathbf{d} = -q_e\mathbf{r}$. All transitions for which the matrix element $\langle f|\mathbf{d}|i\rangle$ vanishes are "forbidden" in the *electric-dipole approximation*: this means that they occur at very low rates, associated to higher multipoles in the field expansion.

The matrix elements of the dipole operator of the one-electron wavefunction are:

$$\langle n_f, l_f, m_{l\,f}|\mathbf{d}|n_i, l_i, m_{l\,i}\rangle = \int \psi^*_{n_f l_f m_{l\,f}}(\mathbf{r})\,\mathbf{d}\,\psi_{n_i l_i m_{l\,i}}(\mathbf{r})\,d^3r. \tag{2.46}$$

This integration is carried out conveniently in polar coordinates: express the dipole operator $\mathbf{d}$ as $-q_e r\ (\sin\theta\ \cos\varphi,\ \sin\theta\ \sin\varphi,\ \cos\theta)$, and observe that

$$\begin{aligned}
r_x &= r\sqrt{\frac{2\pi}{3}}\,\left[Y_{1\,-1}(\theta,\varphi) - Y_{1\,1}(\theta,\varphi)\right]\\
r_y &= r\,i\sqrt{\frac{2\pi}{3}}\,\left[Y_{1\,-1}(\theta,\varphi) + Y_{1\,1}(\theta,\varphi)\right]\\
r_z &= r\sqrt{\frac{4\pi}{3}}\,Y_{1\,0}(\theta,\varphi)
\end{aligned}$$

[the inverse of Eq. (2.18)], so that the squared dipole matrix element is proportional to

$$\begin{aligned}
|\langle f|\mathbf{r}|i\rangle|^2 &= |\langle f|r_x|i\rangle|^2 + \left|\langle f|r_y|i\rangle\right|^2 + |\langle f|r_z|i\rangle|^2\\
&= |\langle f|r|i\rangle|^2\,\frac{2\pi}{3}\Big(|\langle f|Y_{1\,-1} - Y_{1\,1}|i\rangle|^2 + |\langle f|Y_{1\,-1} + Y_{1\,1}|i\rangle|^2\\
&\qquad\qquad + 2|\langle f|Y_{1\,0}|i\rangle|^2\Big)\\
&= |\langle f|r|i\rangle|^2\,\frac{4\pi}{3}\left(|\langle f|Y_{1\,-1}|i\rangle|^2 + |\langle f|Y_{1\,1}|i\rangle|^2 + |\langle f|Y_{1\,0}|i\rangle|^2\right),
\end{aligned}$$

with factored radial and angular matrix elements. Explicitly, in terms of the wavefunctions

$$\begin{aligned}
&\left|\langle n_f, l_f, m_{l\,f}\right|\mathbf{d}\,|n_i, l_i, m_{l\,i}\rangle|^2\\
&\quad= q_e^2\left|\int_0^\infty R_{n_f\,l_f}(r)\,r\,R_{n_i\,l_i}(r)\,r^2\,dr\right|^2\\
&\qquad\times\frac{4\pi}{3}\sum_m\left|\int_0^\pi \sin\theta\,d\theta\int_0^{2\pi} d\varphi\,Y^*_{l_f\,m_f}(\theta,\varphi)Y_{1\,m}(\theta,\varphi)Y_{l_i\,m_i}(\theta,\varphi)\right|^2. \qquad (2.47)
\end{aligned}$$

No radial integral vanishes, because the function $r\ R_{n_i\,l_i}(r)$ has nonzero expansion coefficients on all radial basis wavefunctions of Eq. (2.13). The radial integral can be computed analytically between any initial (n_i, l_i) and final (n_f, l_f): it decreases rapidly when n_i and n_f differ by large amounts, because $R_{n_i\,l_i}(r)$ and $R_{n_f\,l_f}(r)$ take nonnegligible values in remote region, so that the product $R_{n_i\,l_i}(r)\ R_{n_f\,l_f}(r)$ is small everywhere.

Much stricter results are found for the angular part: the integration in the second row of Eq. (2.47) represents the angular overlap of a state with angular momentum $l = l_f$ with the product $Y_{1\,m}\ Y_{l_i\,m_i}$. By applying the rule (B.72) of angular-momentum coupling, this product of objects carrying angular momenta $l = 1$ and

$l = l_i$ can be decomposed into the following allowed values[4] of the total coupled angular momentum:

$$l_f = |l_i - 1|, \; l_i, \; l_i + 1. \tag{2.48}$$

Final states characterized by angular momentum l_f *not* satisfying Eq. (2.48) are guaranteed to make the angular integral vanish. Moreover, the angular integral vanishes also when $l_f = l_i$. The reason is that the parity of $Y_{1\,m}$, like that of **r**, is $(-1)^1 = -1$, while the parity of $Y^*_{l_i\,m_{l\,f}} Y_{l_i\,m_{l\,i}}$ is $(-1)^{2l_i} = 1$. As a result, the parity of the integrated product function $Y^*_{l_i\,m_f} Y_{1\,m} Y_{l_i\,m_i}$ is odd, and therefore its integration over the entire solid angle vanishes.

In summary, in the dipole approximation, nonzero matrix element can occur only for transitions involving states with l changing by exactly unity. The *allowed transitions* have therefore

$$\Delta l = l_f - l_i = \pm 1. \tag{2.49}$$

This equality is the *electric-dipole selection rule* regarding l.

The fact that the dipole operator is associated to an orbital $l = 1$ implies also that its component $m = -1, 0, 1$. Accordingly, the only value of $m_{l\,f}$ for which the angular integral is nonzero, is obtained by adding m to $m_{l\,i}$. From this observation we formulate the m_l-selection rule:

$$\Delta m_l = m_{l\,f} - m_{l\,i} = 0, \; \pm 1. \tag{2.50}$$

Until this point, spin was ignored because the dipole operator is purely spatial: it does nothing to spin. Indeed in the uncoupled basis $|n, l, m_l, m_s\rangle$, see Appendix B.8.1, **d** only affects the spatial degrees of freedom, while for spin it acts as the identity. As a result, we have

$$\Delta s = s_f - s_i \; = \; 0 \tag{2.51}$$

$$\Delta m_s = m_{sf} - m_{si} = 0. \tag{2.52}$$

The first spin selection rule (2.51) is trivial for a one-electron atom, as $s \equiv 1/2$ anyway, but it will become relevant for many-electron atoms.

By analyzing the composition of the coupled states $|n, l, j, m_j\rangle$ in terms of the $|n, l, m_l, m_s\rangle$ uncoupled basis, see Eq. (B.73), one can obtain the following selection rules for the coupled states:

---

[4] In general it can be shown that the angular integral within the absolute value in Eq. (2.47) is proportional to the Clebsch-Gordan coefficient $C^{l_f\,m_f}_{l_i\,m_{l\,i}\;1\,m}$, see Eq. (B.73).

$$\Delta j = j_f - j_i = 0, \pm 1 \qquad (\text{not } 0 \longrightarrow 0) \tag{2.53}$$

$$\Delta m_j = m_{j\,f} - m_{j\,i} = 0, \pm 1. \tag{2.54}$$

After determining what transitions are allowed in the dipole approximation, we should use Eq. (2.45) to estimate typical rates of radiative atomic transitions. By noting that $\mathscr{E}_{if} \approx Z^2 E_{\text{Ha}}$ and observing that the order of magnitude of $|\langle f|\mathbf{d}|i\rangle| \approx q_e a_0/Z$, we can estimate

$$\begin{aligned}\gamma_{if} &= \frac{\mathscr{E}_{if}^3\,|\langle f|\mathbf{d}|i\rangle|^2}{3\pi\,\varepsilon_0\hbar^4c^3} \simeq \frac{Z^4}{\varepsilon_0\hbar^3c^3}E_{\text{Ha}}^2\,\omega_{if}\,q_e^2\,\frac{a_0^2}{Z^2} \simeq \frac{e^2Z^2}{(\hbar c)^3}e^4\omega_{if} = Z^2\left(\frac{e^2}{\hbar c}\right)^3\omega_{if} \\ &= Z^2\alpha^3\omega_{if},\end{aligned} \tag{2.55}$$

where we have used $E_{\text{Ha}}a_0 = e^2$ and $\alpha = e^2/(\hbar c)$. As $\alpha^3 \simeq 10^{-7}$ and in one-electron atoms $\omega_{if} = \mathscr{E}_{if}/\hbar \simeq Z^2\,10^{16}\,\text{Hz}$, we expect typical radiative transition rates of the order $\gamma_{if} \simeq Z^4\,10^9\,\text{s}^{-1}$, i.e. decay times $\gamma_{if}^{-1} \simeq Z^{-4}\,\text{ns}$. The strong ($\mathscr{E}_{if}^3$) dependence of γ_{if} makes this rate faster for more energetic transitions, and slower for low-energy transitions. Much slower transition occur for *dipole-forbidden* transitions, associated to weaker higher-order couplings to the electromagnetic field (magnetic dipole, electric quadrupole…).

In a gaseous sample, other *nonradiative* transitions can also occur due to collisions with other atoms and/or with the vessel walls. Nonradiative mechanisms can dominate the decay of long-lived *metastable* states, i.e. those lacking fast dipole-allowed decay transitions.

2.1.10 Spectra in a Magnetic Field

We conclude this Section with a brief analysis of atomic spectra in the condition where a maximum of information can be extracted from them, namely when the atomic sample is immersed in a uniform magnetic field. In these conditions, the atom, depending on the component of its magnetic moment along the field direction, acquires a little extra energy which can then be detected by spectroscopy.

In the presence of both orbital and spin angular momenta, the total atomic magnetic moment is the vector sum

$$\boldsymbol{\mu} = \boldsymbol{\mu}_l + \boldsymbol{\mu}_s = -\mu_{\text{B}}\,\frac{g_l\mathbf{L} + g_s\mathbf{S}}{\hbar} \simeq -\mu_{\text{B}}\,\frac{\mathbf{L} + 2\mathbf{S}}{\hbar}. \tag{2.56}$$

By studying the matrix elements of the $\boldsymbol{\mu}$ operator, one can evaluate the magnetic properties of an atom, where both spin and orbital magnetic moments come into play. Following Eqs. (2.24) and (2.56), the coupling with the external magnetic field can be expressed as:

$$H_{\text{magn}} = -\mathbf{B}\cdot\boldsymbol{\mu} = \mu_{\text{B}}\,\mathbf{B}\cdot\frac{\mathbf{L}+2\mathbf{S}}{\hbar} = \mu_{\text{B}}\,B_z\,\frac{L_z+2S_z}{\hbar}. \tag{2.57}$$

This operator is diagonal in the uncoupled $|l, s, m_l, m_s\rangle$ basis. Remember however (Sect. 2.1.7) that $H_{\text{s-o}}$ is not diagonal in that basis, but rather in the coupled basis $|l,\ s,\ j,\ m_j\rangle$. In fact, $H_{\text{s-o}}$ and H_{magn} cannot be diagonalized simultaneously, as they do not commute. To obtain the eigenenergies and eigenkets one must diagonalize the total operator $H_{\text{magn}} + H_{\text{s-o}}$, within each $(2l+1)\cdot(2s+1)$-dimensional subspace at fixed n, l, and s. This diagonalization is not especially complicated, but it is perhaps more instructive to understand in detail the two limiting cases where either characteristic energy scale $\mu_{\text{B}}\,|\mathbf{B}|$ or ξ dominates.

The simplest limit ($\mu_{\text{B}}|\mathbf{B}| \gg \xi$), of magnetic field energy $\mu_{\text{B}}|\mathbf{B}|$ much larger than the spin-orbit energy ξ, occurs for sufficiently large field intensity. The "sufficiently large" threshold depends on the considered atom and level: for hydrogen 2p, the strong-field limit is reached for $|\mathbf{B}| \gg 0.5$ T, while for He^+ 2p it takes a magnetic field as large as $|\mathbf{B}| \gg 8$ T, due to the Z^4-dependence of the spin-orbit energy, Eq. (2.35). In this limit of very strong field, the coupled basis is not especially good, as full rotational invariance of the atom is badly broken. The uncoupled basis $|l, s, m_l, m_s\rangle$ works fine instead: spin and orbital moments align relative to the field with separate energy contributions depending on their different g-factors. In this basis, the dominating interaction H_{magn} is diagonal: if we neglect the smaller $H_{\text{s-o}}$, the magnetic energy levels are simply

$$E_{\text{magn}}(m_l, m_s) \simeq \langle m_l,\ m_s|H_{\text{magn}}|m_l,\ m_s\rangle = \mu_{\text{B}}\,B_z(m_l+2m_s) \tag{2.58}$$

(*Paschen-Back limit*). $H_{\text{s-o}}$ corrections may be added perturbatively.

In the (more common) opposite weak-field limit ($|\mathbf{B}| \ll \xi/\mu_{\text{B}}$), spherical symmetry is perturbed only weakly. The states $|l,\ s,\ j,\ m_j\rangle$ in the coupled basis are exact eigenstates of $H_{\text{s-o}}$ and approximate eigenstates of $H_{\text{magn}} + H_{\text{s-o}}$. Following the results of Appendix B.8.2, to first order in $\mu_{\text{B}}|\mathbf{B}|/\xi$, the energy contribution of H_{magn} is given by Eq. (B.75):

$$E_{\text{magn}}(m_j) \simeq \langle j, m_j|H_{\text{magn}}|j, m_j\rangle = \langle j, m_j| - \mu_z B_z|j, m_j\rangle = g_j\,\mu_{\text{B}}\,B_z m_j \tag{2.59}$$

(*Zeeman limit*), where g_j is the Landé g-factor, Eq. (B.78).

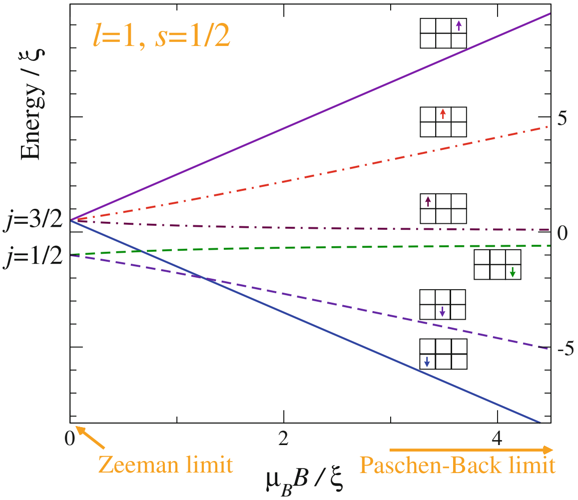


Fig. 2.12 Spin-orbit and magnetic splitting of a ^{2}P multiplet. With the shorthand $b = \mu_B B/\xi$, the expression for the $m_j = \pm 3/2$ energies (*solid lines*) is simply $(1/2 \pm 2b)\,\xi$. The energies of the four other levels are $1/4(-1 \pm 2b + d)\xi$ (*dot-dashed lines*) and $1/4(-1 \pm 2b - d)\xi$ (*dashed lines*), where $d = \sqrt{9 + 4b(1 + b)}$

In the intermediate-field regime $\mu_B|\mathbf{B}| \simeq \xi$ neither basis is appropriate and neither of expressions (2.58) and (2.59) is accurate. Figure 2.12 displays the exact pattern of splittings of the six ^{2}P states under the action of a magnetic field, changing from the Zeeman (weak field) to the Paschen-Back (strong-field) limit. The initial slopes of the energy curves at $B \to 0$, divided by the relevant m_j, measure the values of the Landé g_j.

The experimentally observed spectra confirm the theory outlined here. For H, provided that a sufficiently strong magnetic field is applied, the Paschen-Bach is relatively straightforward to observe as a triplication of all lines. If very high spectral resolution can be achieved, also the weak-field Zeeman splitting of the H lines shown in the conceptual scheme of Fig. 2.13 can be detected. Such Zeeman effect is called "anomalous" since the lines are spaced irregularly. In fact, in atomic physics, such irregular splittings are rather the rule than an anomaly: only selected spin-orbit-free $S = 0$ lines in many-electron atoms happen to show "regular" Zeeman splittings (see Fig. 2.30 in Sect. 2.2.8.4).

2.2 Many-Electron Atoms

2.2.1 *Identical Particles*

The concept of indistinguishable particles is central to all the physics of matter where two or more electrons are involved. Note that in classical mechanics each particle is labeled by its own position and momentum: one could in principle track individual trajectories along the motion, and thus tell identical particles i and j apart at any time.

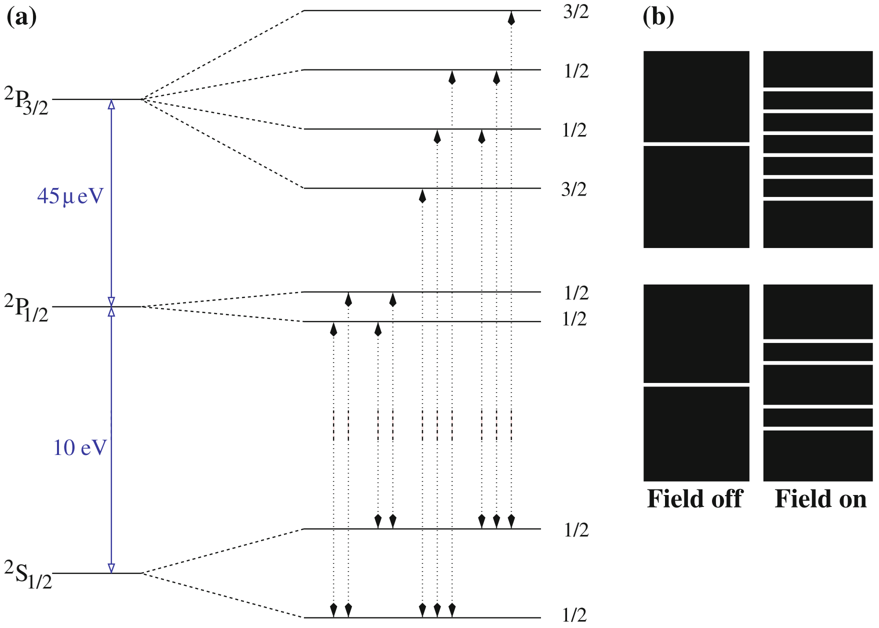


Fig. 2.13 **a** A scheme of the Zeeman-split 1s⟷2p lowest Lyman line of H. Splittings of levels of successive m_j equal the magnetic energy $\mu_B B$ times the appropriate g-factor: $g_j(^2P_{3/2}) = \frac{4}{3}$, $g_j(^2P_{1/2}) = \frac{2}{3}$, $g_j(^2S_{1/2}) = 2$. **b** The line spectrum: the splitting of the nearest *lines* equals $\frac{2}{3}\mu_B B$

In QM, indistinguishable particles are such at the deepest level. There is no way, even as a matter of principle, to ever tell e.g. two electrons apart, as they both are present at all points of space, with a certain probability amplitude. QM implements perfect indistinguishableness through symmetry: any many-particles ket has a definite symmetry "character" for the permutation operator P_{ij} swapping the ith and jth identical particles. As this permutation symmetry is a discrete symmetry which, if applied twice, leads back to the initial state, the eigenvalues of P_{ij} can only be $+1$ or -1.

The particles for whose swap the overall system ket $|a\rangle$ is *symmetric* are called *bosons*. The eigenvalue of P_{ij} is $+1$, i.e. $P_{ij}|a\rangle = |a\rangle$.

The particles for whose swap the overall ket $|a\rangle$ is *antisymmetric* are called *fermions*. Here the eigenvalue of P_{ij} is -1, i.e. $P_{ij}|a\rangle = -|a\rangle$.

All elementary particles of "matter" (electrons, protons, neutrons[5]…) are examples of fermions, all of spin 1/2. Elementary bosons are carriers of "interactions". For example, photons (spin 1) are the quanta of the electromagnetic field, see Sect. 4.3.2.2. A simple general rule connects the spin of a particle kind to its permutational symmetry: *integer-spin particles are bosons, half odd-integer particles are fermions.*

---

[5] Contrary to electrons, protons and neutrons are not quite elementary, since they are bound states of triplets of elementary spin-1/2 quarks, and they follow the rule for composite particles sketched below.

A collection of bosons and fermions lumped together can behave as a single point-like particle. This occurs when the internal dynamics is associated to high excitation energy, so that the lump remains in its ground state (possibly degenerate according to the projection of the lump's total angular momentum). It is important to know the eigenvalue for the permutation of two such *identical* lumps (*composite bosons/fermions*). This is simply answered by counting the number of (-1)'s generated by the permutations of pairs of identical fermions. For example, hydrogen atoms in the same hyperfine state are identical bosons, since a -1 is generated by the permutation of the two electrons and a second -1 is generated by the permutation of the two protons, in total $(-1) \cdot (-1) = +1$. Likewise, the ^{13}C isotope of carbon is a fermion (6 protons + 7 neutrons + 6 electrons), while ^{238}U is a boson (92 protons + 146 neutrons + 92 electrons). Similarly, identical nuclei are bosons or fermions according to whether they contain an even or odd number of nucleons (protons plus neutrons); for example the deuteron nucleus D^+, a bound proton and neutron in a $i = 1$ nuclear-spin state, is a boson. Due to the angular-momentum composition rule (B.72), all composite particles fulfill the general spin rule.

The permutational symmetry is crucial to understand the dynamics of many electrons and, in particular, the structure of many-electron atoms. As discussed in greater detail below, antisymmetry obliges N electrons to span N quantum states, thus effectively avoiding one another. In practice, the geometrical constraint of antisymmetry is often more effective than the dynamical electron–electron repulsion (1.6) in keeping electrons apart.

Without permutational antisymmetry all electrons would occupy the same 1s shell in the atomic ground state. If that happened (as it could if electrons were distinguishable particles—or bosons—rather then fermions), then the atomic ionization energies should increase roughly as Z^2 and the size of atoms should decrease roughly as Z^{-1}. In stark contrast, relatively mild non-monotonic Z dependence of both these properties (Figs. 2.14 and 2.15) are observed: in particular, the first ionization energy and the atomic size exhibit a weak general tendency to respectively decrease and increase with Z.

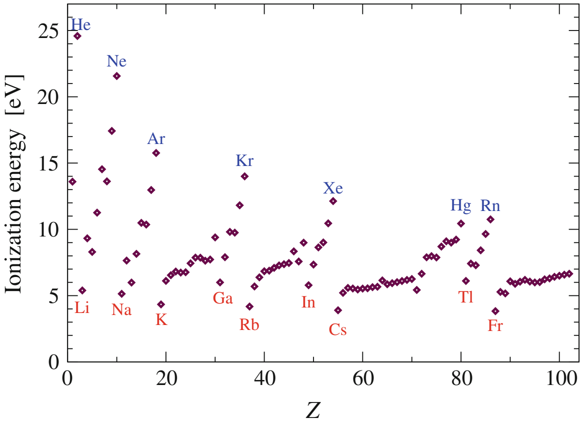


Fig. 2.14 As a function of the atomic number Z, the first-ionization energy of neutral atoms ($N = Z$), i.e. the minimum energy required to remove one electron from the atom

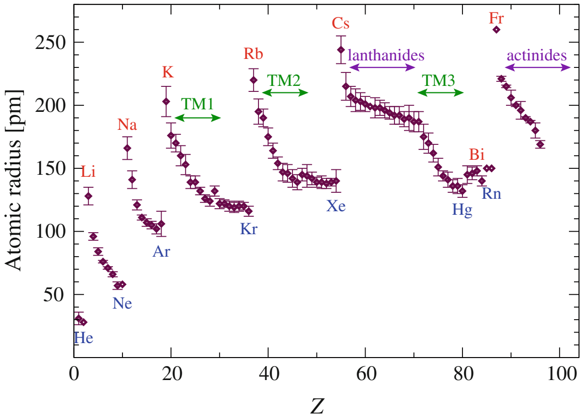


Fig. 2.15 Empirical atomic radii, as a function of Z. The vertical bars represent the standard deviation. The labels *TM1*, *TM2* and *TM3* mark the ranges of the *first*, *second* and *third* transition series, respectively (Data from Ref. [19])

2.2.2 The Independent-Particles Approximation

The exact solution of the Schrödinger equation associated to Hamiltonian (1.1) requires the determination of a N-electron wavefunction. For increasing N, this becomes soon a formidable task, because the N-electron wavefunction describes the correlated motion of all the N electrons, and thus depends intricately on all position and spin coordinates of the N electrons. The amount of information carried by a generic N-electron wavefunction is exponentially large with N: there is no way to store (let alone compute!) the full wavefunction for the ground state of many interacting electrons.

Most approximate methods on the market exploit the observation that a basis of the Hilbert space of N-particle states can be built as a product of single-particle basis states. Consider the complete set of orthonormal states $\{|\alpha\rangle\}$ for a single particle, where α takes all possible values allowed to a full set of quantum numbers, e.g. n, l, m_l, m_s in the example of an electron in an atom. Then the tensor product

$$|\alpha_1, \alpha_2, \ldots, \alpha_N\rangle = |\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \cdots |\alpha_N\rangle, \tag{2.60}$$

realizes a basis for N particles when all possible choices of $\alpha_1, \alpha_2, \ldots, \alpha_N$ are explored. For indistinguishable particles, the correct permutational symmetry is imposed to the product state (2.60) by taking the properly symmetrized linear combination

$$\begin{aligned} |\alpha_1, \alpha_2, \ldots, \alpha_N\rangle^{S/A} &= \frac{1}{\sqrt{\mathcal{N}_P}} \sum_P (\pm 1)^{\{P\}} P |\alpha_1, \alpha_2, \ldots, \alpha_N\rangle \\ &= \frac{1}{\sqrt{\mathcal{N}_P}} \sum_P (\pm 1)^{\{P\}} |\alpha_{P_1}, \alpha_{P_2}, \ldots, \alpha_{P_N}\rangle. \end{aligned} \tag{2.61}$$

Here P indicates a generic permutation of the N states α_j, and the sum extends over all $\mathscr{N}_P$ permutations; for example, $\mathscr{N}_P = N!$ if the states α_j happen to be all different. $\{P\}$ in the exponent indicates the parity of the permutation P, i.e. the number of pair swaps composing P. The fully symmetrized basis state $|\alpha_1, \alpha_2, \ldots, \alpha_N\rangle^S$ realizes the correct permutational symmetry of N bosons; the antisymmetric combination $|\alpha_1, \alpha_2, \ldots, \alpha_N\rangle^A$ involving nontrivial (-1) signs can play the role of basis state for N fermions. For bosons, no restriction applies to the quantum numbers α_j: any number of them may coincide. Instead, *for fermions, all quantum numbers must necessarily be different*. If two were equal, say $\alpha_i = \alpha_j$, in the sum (2.61), the kets $|\alpha_{P_1}, \ldots, \alpha_{P_i}, \ldots, \alpha_{P_j}, \ldots, \alpha_{P_N}\rangle$ and $|\alpha_{P_1}, \ldots, \alpha_{P_j}, \ldots, \alpha_{P_i}, \ldots, \alpha_{P_N}\rangle$ would be equal, but with opposite parity phase factor $(-1)^{\{P\}}$, so that they all cancel in pairs in the sum, and the total ket $|\alpha_1, \ldots, \alpha_i, \ldots, \alpha_j, \ldots, \alpha_N\rangle^A$ vanishes. As a result, the product basis kets for N fermions are characterized by N different quantum numbers: this property expresses the *Pauli exclusion principle*, according to which *two identical fermions can never occupy the same quantum state*.

The wavefunction representation of a N-fermion basis ket is obtained starting from the eigenket of position and spin of the jth fermion, shorthanded to $|w_j\rangle = |\mathbf{r}_j, \sigma_j\rangle$. The corresponding N-fermions product converts to a bra as $\langle w_1, w_2, \ldots w_N|$, and is then to be properly antisymmetrized. The wavefunction associated to Eq. (2.61) is then

$$\begin{aligned}\Psi_{\alpha_1,\ldots,\alpha_N}(w_1, \ldots, w_N) &= {}^A\langle w_1, w_2, \ldots, w_N|\alpha_1, \alpha_2, \ldots, \alpha_N\rangle^A \\ &= \frac{1}{N!}\sum_{P'}(-1)^{\{P'\}}\sum_P(-1)^{\{P\}}\langle w_{P'_1}, \ldots, w_{P'_N}|\alpha_{P_1}, \ldots, \alpha_{P_N}\rangle \\ &= \frac{1}{N!}\sum_{P'}(-1)^{\{P'\}}\sum_P(-1)^{\{P\}}\psi_{\alpha_{P_1}}(w_{P'_1})\ldots\,\psi_{\alpha_{P_N}}(w_{P'_N}) \\ &= \sum_P(-1)^{\{P\}}\psi_{\alpha_{P_1}}(w_1)\ldots\,\psi_{\alpha_{P_N}}(w_N),\end{aligned} \tag{2.62}$$

where the sum over permutations of the variables w_j generates $N!$ identical copies of the same wavefunction, and is therefore eliminated. The sum in the last expression is the determinant of the matrix whose elements are $\psi_{\alpha_i}(w_j)$:

$$\Psi_{\alpha_1,\ldots,\alpha_N}(w_1, \ldots, w_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix}\psi_{\alpha_1}(w_1) & \psi_{\alpha_1}(w_2) & \cdots & \psi_{\alpha_1}(w_N) \\ \psi_{\alpha_2}(w_1) & \psi_{\alpha_2}(w_2) & \cdots & \psi_{\alpha_2}(w_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{\alpha_N}(w_1) & \psi_{\alpha_N}(w_2) & \cdots & \psi_{\alpha_N}(w_N)\end{vmatrix}. \tag{2.63}$$

This basis wavefunction is called a *Slater determinant*.[6]

---

[6] In Eq. (2.63) we introduce an extra normalization factor $(N!)^{-1/2}$. This standard factor allows one to carry out integration (e.g. for wavefunction normalization) over an unrestricted domain of all w_j variables, instead of the appropriate "hyper-triangle" $w_1 > w_2 > \cdots > w_N$.

An (anti)symmetrized product ket (2.61) contains an amount of information only directly, rather than exponentially, proportional to the number N of identical particles, and is thus substantially simpler than a general boson/fermion ket. Despite their simplicity, the kets of Eq. (2.61) constitute a basis of the Hilbert space of the proper symmetry: any physical N-boson/N-fermion ket $|a^{B/F}\rangle$, e.g. an exact energy eigenstate of N interacting identical particles, can be expressed as a linear combination of the factorized basis kets:

$$|a^{B/F}\rangle = \sum_{\alpha_1,\alpha_2,\ldots,\alpha_N} c^a_{\alpha_1,\alpha_2,\ldots,\alpha_N} |\alpha_1, \alpha_2, \ldots, \alpha_N\rangle^{S/A}, \tag{2.64}$$

where $c^a_{\alpha_1,\alpha_2,\ldots,\alpha_N}$ are the complex coefficients defining the linear combination. The number of these coefficients grows exponentially with N: the amount of information of the correlated state $|a^{B/F}\rangle$ is now encoded in its expansion coefficients $c^a_{\alpha_1,\alpha_2,\ldots,\alpha_N}$.

Several approximate methods of solution of the Schrödinger problem for many-electron systems replace the (lowest) exact eigenstate with *one* (as smart as possible) basis state $|\alpha_1, \alpha_2, \ldots, \alpha_N\rangle^A$ constructed with single-particle states $|\alpha_1\rangle$, $|\alpha_2\rangle$,… solutions of some appropriate single-electron Hamiltonian. This scheme is named *independent-particles approximation*.

For atoms, the simplest approach along this line consists in neglecting the electron–electron Coulomb interaction V_{ee} of Eq. (1.6) altogether. In this approximation, the Schrödinger problem for the N electrons factorizes exactly: each electron moves independently of the others in the field V_{ne} of the nucleus of charge Zq_e. Neglecting relativistic corrections, the single-electron eigenstates $|\alpha_j\rangle$ are represented by hydrogenic wavefunctions of the kind (2.27), as defined in Eqs. (2.11)–(2.13), with the appropriate Z.[7] In this atomic context α_j stands for a set of quantum numbers $n_j, l_j, m_{l\,j}, m_{s\,j}$. Choose a list of N *different* α_j: the generic N-electron eigenstate is $|\alpha_1, \alpha_2, \ldots \alpha_N\rangle^A$, as in Eq. (2.61). The example of a possible $N = 4$-electron state is

$$|1, 0, 0, \uparrow,\ 3, 1, 1, \uparrow,\ 3, 1, 0, \uparrow,\ 3, 1, -1, \downarrow\rangle^A. \tag{2.65}$$

The standard spectroscopic notation $1\mathrm{s}3\mathrm{p}^3$ for this state lists the occupied single-particle orbitals, with the corresponding electron numbers as exponents: all information about the m_l's and m_s's is omitted. In contrast, the box notation $1\mathrm{s}\begin{array}{|c|}\hline \uparrow \\ \hline \\ \hline\end{array}\ 3\mathrm{p}\begin{array}{|c|c|c|}\hline & \uparrow & \uparrow \\ \hline \downarrow & & \\ \hline\end{array}$ contains all details.

The total energy E^{tot} of an atomic state is minus the work needed to decompose the atom from that given bound state to an isolated nucleus plus the N individual electrons at rest at infinite reciprocal distance. The total energy of non-interacting electrons is simply the sum of their single-electron energies.

---

[7] We neglect all reduced-mass effects, assuming an infinite nuclear mass. The finiteness of the nuclear mass could introduce tiny correlations in the electronic motions, which would add to those of Coulombic origin.

The eigenstate (2.65) is not the one with the lowest possible energy. The ground state can be obtained by minimizing the energy of each single-particle orbital, *without violating the Pauli principle*. For $N = 4$, the lowest energy is realized by any of $1s\begin{array}{|c|}\hline \uparrow \\ \hline \downarrow \\ \hline \end{array}\,2s\begin{array}{|c|}\hline \uparrow \\ \hline \downarrow \\ \hline \end{array}$, $1s\begin{array}{|c|}\hline \uparrow \\ \hline \downarrow \\ \hline \end{array}\,2s\begin{array}{|c|}\hline \uparrow \\ \hline \\ \hline \end{array}\,2p\begin{array}{|c|c|c|}\hline & & \uparrow \\ \hline & & \\ \hline \end{array}$ plus 11 similar states, $1s\begin{array}{|c|}\hline \uparrow \\ \hline \downarrow \\ \hline \end{array}\,2p\begin{array}{|c|c|c|}\hline & & \uparrow \\ \hline & \downarrow & \\ \hline \end{array}$ plus 14 similar states: a total of 1 ($1s^2 2s^2$) +12 ($1s^2 2s 2p$) +15 ($1s^2 2p^2$) = 28 individual states, all with energy

$$E^{\text{tot}}_{1s^2 2s^2} = 2\mathscr{E}_1 + 2\mathscr{E}_2 = -2\,\frac{E_{\text{Ha}}}{2}\left(\frac{Z^2}{1^2} + \frac{Z^2}{2^2}\right) = -\frac{5}{4}\,Z^2\,E_{\text{Ha}}. \tag{2.66}$$

The excited state (2.65) has significantly higher total energy $E^{\text{tot}}_{1s3p^3} = -\frac{2}{3}\,Z^2\,E_{\text{Ha}}$.

Real electrons do interact with each other. In a neutral atom, the electron–electron repulsion (1.4) is of the same order of magnitude as the attraction (1.6) to the nucleus, which means that the brutal simplification of a complete neglect of V_{ee} is a very poor approximation, doomed to yield unphysical predictions. Since the maximum occupation of the nth hydrogenic level $2n^2$ grows rapidly, the minimum energy required to remove an electron from a neutral atom (the first ionization energy) increases with Z only marginally more slowly than Z^2, at variance with experiment (Fig. 2.14). Moreover, in this model any atom (regardless of Z) would be able to accept any number N of electrons, always forming bound states. Experimentally, however, only certain atoms can form negatively-charged ionic bound states, but never with more than 1 extra charge ($N \leq Z + 1$). The main reason for the failures of this model is illustrated in Fig. 2.16: in reality while an electron is removed from the atom, it does not feel the bare nuclear attraction $-Ze^2/r$ but rather, due to electron–electron repulsion and according to the divergence law of electromagnetism, the substantially weaker combined effect of the nuclear charge and that of the other $N-1$ electrons, $V(r) \simeq (-Z + N - 1)\,e^2/r$. This phenomenon of *screening* decreases the ionization energy substantially, and must be included for a fair description of the atomic wavefunction.

The following section sketches a simple but instructive theory for the simplest many-electron atom, He ($Z = N = 2$), for which significant insight is obtained by treating V_{ee} as a perturbation to the uncorrelated electron states. Perturbative methods fail for $N \geq 3$: Sect. 2.2.4 sketches a more systematic method for a much improved independent-electron method, which includes screening.

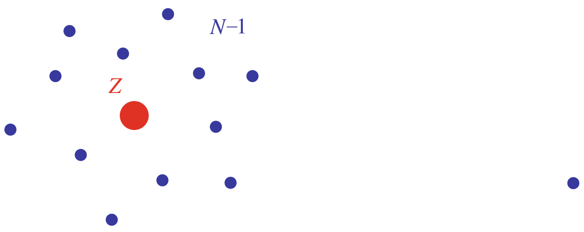


Fig. 2.16 When one electron moves away from an atom/ion, the nucleus of charge Zq_e and the remaining $(N-1)$ electrons attract it as if effectively they were a point charge $(Z - N + 1)\,q_e$

2.2.3 The 2-Electron Atom

In the simplest nontrivial case of $N = 2$ electrons (relevant for the He atom, and for the Li^+, Be^{2+}… ions), the independent-electron wavefunction reads

$$\Psi_{\alpha_1,\alpha_2}(w_1, w_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_{\alpha_1}(w_1) & \psi_{\alpha_1}(w_2) \\ \psi_{\alpha_2}(w_1) & \psi_{\alpha_2}(w_2) \end{vmatrix} = \frac{\psi_{\alpha_1}(w_1)\psi_{\alpha_2}(w_2) - \psi_{\alpha_1}(w_2)\psi_{\alpha_2}(w_1)}{\sqrt{2}}. \tag{2.67}$$

In the cases where the orbital quantum numbers coincide ($n_1 = n_2$, $l_1 = l_2$, $m_{l\,1} = m_{l\,2}$), the orbital part can be factorized in the Slater determinant. Only spin remains in the antisymmetric combination:

$$\Psi_{n,l,m_l,\uparrow,\; n,l,m_l,\downarrow}(w_1, w_2) = \psi_{n,l,m_l}(\mathbf{r}_1)\, \psi_{n,l,m_l}(\mathbf{r}_2)\, \frac{1}{\sqrt{2}} \begin{vmatrix} \chi_\uparrow(\sigma_1) & \chi_\uparrow(\sigma_2) \\ \chi_\downarrow(\sigma_1) & \chi_\downarrow(\sigma_2) \end{vmatrix}. \tag{2.68}$$

Here the Slater determinant is simply $\chi_\uparrow(\sigma_1)\,\chi_\downarrow(\sigma_2) - \chi_\uparrow(\sigma_2)\,\chi_\downarrow(\sigma_1)$. This combination of the two electron spin wavefunctions is an eigenstate of the square modulus of the total spin $\mathbf{S} = \mathbf{s}_1 + \mathbf{s}_2$, with null eigenvalue $|\mathbf{S}|^2 = S(S+1)\hbar^2 = 0$. Eigenstates of $|\mathbf{S}|^2$, like the one of Eq. (2.68), are useful because the matrix elements of the (hitherto neglected) Coulomb repulsion between states of different S vanish since V_{ee} is an orbital operator, which does not act on spin. The other $S = 0$ states (*spin singlets*), those involving two different sets of orbital quantum numbers, are:

$$\begin{aligned} &\Psi^{S=0}_{n_1,l_1,m_{l\,1},\; n_2,l_2,m_{l\,2}}(w_1, w_2) \\ &\quad = \frac{\psi_{n_1,l_1,m_{l\,1}}(\mathbf{r}_1)\, \psi_{n_2,l_2,m_{l\,2}}(\mathbf{r}_2) + \psi_{n_1,l_1,m_{l\,1}}(\mathbf{r}_2)\, \psi_{n_2,l_2,m_{l\,2}}(\mathbf{r}_1)}{\sqrt{2}} \; \frac{1}{\sqrt{2}} \begin{vmatrix} \chi_\uparrow(\sigma_1) & \chi_\uparrow(\sigma_2) \\ \chi_\downarrow(\sigma_1) & \chi_\downarrow(\sigma_2) \end{vmatrix}. \end{aligned} \tag{2.69}$$

Note that the $S = 0$ states of Eq. (2.69) are *not* single Slater determinants of the type of Eq. (2.67). The singlet states are characterized by an orbital part of the wavefunction which is symmetric under permutation P_{12}, with the spin part taking care of the required antisymmetry.

The singlet states (2.68) and (2.69) are $S = 0$ eigenstates of $|\mathbf{S}|^2$. The other value of S allowed by Eq. (B.72) is $S = 1$. The spin part of the wavefunctions of the $S = 1$ *spin-triplet* states is any of:

$$\begin{aligned} \mathcal{X}^{S=1,M_S=1}(\sigma_1, \sigma_2) &= \chi_\uparrow(\sigma_1)\, \chi_\uparrow(\sigma_2) && [M_S = 1] \\ \mathcal{X}^{S=1,M_S=0}(\sigma_1, \sigma_2) &= \tfrac{1}{\sqrt{2}}[\chi_\uparrow(\sigma_1)\, \chi_\downarrow(\sigma_2) + \chi_\uparrow(\sigma_2)\, \chi_\downarrow(\sigma_1)] && [M_S = 0] \quad , \\ \mathcal{X}^{S=1,M_S=-1}(\sigma_1, \sigma_2) &= \chi_\downarrow(\sigma_1)\, \chi_\downarrow(\sigma_2) && [M_S = -1] \end{aligned} \tag{2.70}$$

Table 2.1 The characters of the fixed-total-spin-S two-electron basis states

	$S = 0$—spin-singlet states	$S = 1$—spin-triplet states
Orbital quantum number	Any	$(n_1, l_1, m_{l\,1}) \neq (n_2, l_2, m_{l\,2})$
Orbital wavefunction	Symmetric	Antisymmetric
Spin quantum numbers	↑ and ↓ (different)	Any
Spin wavefunction	Antisymmetric	Symmetric

which are all symmetric for P_{12}. Therefore the orbital part takes care of antisymmetry:

$$\begin{aligned}\Psi^{S=1,M_S}_{n_1,l_1,m_{l\,1},\ n_2,l_2,m_{l\,2}}&(w_1, w_2)\\ &= \frac{\psi_{n_1,l_1,m_{l\,1}}(\mathbf{r}_1)\,\psi_{n_2,l_2,m_{l\,2}}(\mathbf{r}_2) - \psi_{n_1,l_1,m_{l\,1}}(\mathbf{r}_2)\,\psi_{n_2,l_2,m_{l\,2}}(\mathbf{r}_1)}{\sqrt{2}}\ \mathscr{X}^{S=1,M_S}(\sigma_1, \sigma_2).\end{aligned} \tag{2.71}$$

In these states, at least one of the orbital quantum numbers for the two electrons needs to be different: $(n_1, l_1, m_{l\,1}) \neq (n_2, l_2, m_{l\,2})$, or else the wavefunction vanishes.

Table 2.1 summarizes the basic properties of the singlet ($S = 0$) and triplet ($S = 1$) basis states $|n_1, l_1, m_{l\,1},\ n_2, l_2, m_{l\,2},\ S, M_S\rangle$ for two electrons. These spin-symmetrized states are convenient 0th-order states for a perturbation theory in V_{ee}, see Appendix B.39. They are eigenstates of $T_e + V_{ne}$: like in the example of Eq. (2.66), the "unperturbed" energies

$$E^{\text{tot}\,(0)}_{n_1,n_2} = -\frac{E_{\text{Ha}}}{2}\left(\frac{Z^2}{n_1^2} + \frac{Z^2}{n_2^2}\right) = -2\,E_{\text{Ha}}\left(\frac{1}{n_1^2} + \frac{1}{n_2^2}\right), \tag{2.72}$$

where the last expression refers to He ($Z = 2$). The ground state $n_1 = n_2 = 1$ (in spectroscopic notation $1s^2$) is necessarily a spin singlet and has energy $E^{\text{tot}\,(0)}_{1,1} = -4\,E_{\text{Ha}} \simeq -109\,\text{eV}$.

As a next step, following standard perturbation theory (Appendix B.9) the effect of V_{ee} is accounted for by evaluating its diagonal matrix elements over the zeroth order eigenkets. Equation (B.85) provides the first-order (additive) correction $E^{\text{tot}\,(1)} = \langle V_{ee}\rangle$ to the eigenenergies:

$$\begin{aligned}E^{\text{tot}\,(1)}_{n_1,l_1,m_{l\,1},\ n_2,l_2,m_{l\,2},\ S,M_S}&\\ = \langle n_1, l_1, m_{l\,1},\ n_2, l_2, m_{l\,2},\ S, M_S&|V_{ee}|n_1, l_1, m_{l\,1},\ n_2, l_2, m_{l\,2},\ S, M_S\rangle.\end{aligned} \tag{2.73}$$

Since the electron–electron repulsion is positive, this correction is *always positive*. The detailed calculation of these Coulomb integrals is a rather intricate mathematical exercise, with instructive qualitative outcomes. The largest of these $E^{\text{tot}\,(1)}$

corrections occurs for the most localized wavefunction, the one where the two electrons stay very close together, both in the 1s level: the ground state $1s^2$. The average inter-electron distance is approximately a_0, thus the Coulomb integral $E^{\text{tot}\,(1)}_{1,0,0,\ 1,0,0,\ 0,0}$ is of order $\sim e^2/a_0 = E_{\text{Ha}}$. The precise value obtained from integration [5] is $\frac{5}{4}\,E_{\text{Ha}} \simeq 34\,\text{eV}$. This correction brings the estimated ground-state energy of He to $E^{\text{tot}\,(0)}_{1,1} + E^{\text{tot}\,(1)}_{1,0,0,\ 1,0,0,\ 0,0} = -74.8\,\text{eV}$, in fair agreement with the experimental value $-79.00\,\text{eV}$ (minus the sum of the first and second ionization energies of He).

For the optically most relevant states, those with one electron sitting in 1s and the other in an excited state $n_2[l_2]$, the first-order correction accounts for several experimental observations:

- All Coulomb integrals are smaller than the one for the ground state, and tend to *decrease for increasing* n_2: the Coulomb correction become less and less important as the electrons move apart from each other.
- The Coulomb integrals at given n_2 depend weakly on l_2: they usually *increase for increasing* l_2. By looking at the hydrogenic radial distributions Fig. 2.7, one can note that indeed the electrons, on average, sit slightly closer when l_2 is larger. This is an important novelty, since it breaks the H-atom l-degeneracy of the shells, putting $\mathscr{E}_{n\text{s}} < \mathscr{E}_{n\text{p}} < \mathscr{E}_{n\text{d}} < \cdots$, in accord to experimental finding (Fig. 2.17).

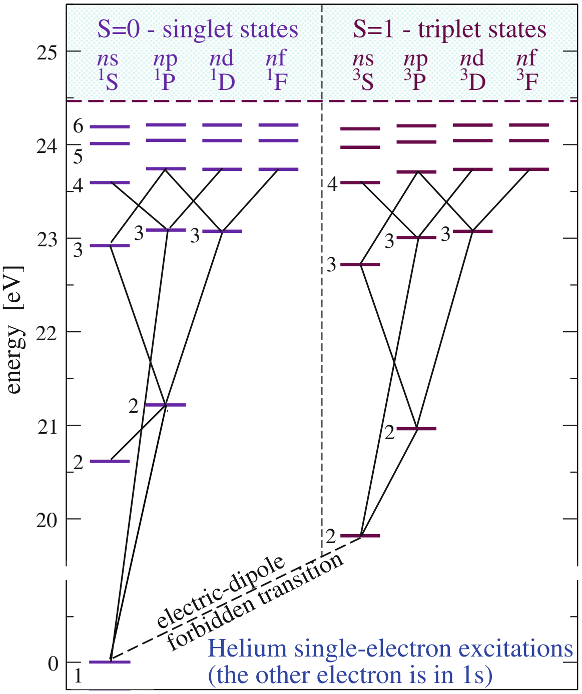


Fig. 2.17 Energy levels of atomic He, with several optical transitions. While one electron remains in the $n_1 = 1$ single-electron state, the quantum numbers $n = n_2$ and l_2 refer to the other (excited) electron. Note that each triplet state sits systematically lower than the corresponding singlet. The energy zero is the ground state $1s^2$, and the *horizontal dashed line* at 24.59 eV marks the first-ionization threshold. (Inspired by Fig. 4.4 of Ref. [3]; data from Ref. [20])

- The Coulomb integrals depend on S, clearly not through the spin wavefunction which has nothing to do with the purely spatial operator V_{ee}, but through the different electron–electron correlation in the P_{12}-symmetric ($S = 0$) or P_{12}-antisymmetric ($S = 1$) spatial wavefunction. In particular, Eq. (2.71) shows that the triplet wavefunction $\Psi^{S=1,M_S}_{n_1,l_1,m_{l\,1},\; n_2,l_2,m_{l\,2}}(w_1, w_2)$ vanishes for $\mathbf{r}_1 \to \mathbf{r}_2$. On the contrary, the singlet wavefunction $\Psi^{S=0}_{n_1,l_1,m_{l\,1},\; n_2,l_2,m_{l\,2}}(w_1, w_2)$ is finite at $\mathbf{r}_1 = \mathbf{r}_2$. Therefore, on average, the electrons in a spin-triplet state avoid each other more effectively than in the spin-singlet state with the same orbital quantum numbers.[8] Indeed, Coulomb integrals are systematically smaller for $S = 1$ than for $S = 0$ states, as explicit evaluation of the integral (2.73) shows. This result accounts for the experimental observation that each triplet state always lies below the corresponding singlet (Fig. 2.17). This kind of Coulomb splitting between states which differ uniquely for their total spin, here $S = 0$ or 1, is called *exchange splitting*.

The perturbative approach presented here is useful mostly as a conceptual tool, to understand qualitative trends, and general concepts such as those listed above. Perturbation theory is relatively successful for the 2-electron atom, but for $N > 2$ electrons the repulsion that a given electron experiences from the other $N - 1$ electrons is comparable to the attraction generated by the nucleus, and any attempt to treat it as a small perturbation fails. A better approximate approach, based on a mean-field self-consistent evaluation of the electron–electron repulsion, yields fair quantitative accuracy for any N and is commonly used to date. The reliability of this and similar self-consistent field methods have made them standard tools for understanding experiments and making predictions of atomic properties of matter from first principles.

2.2.4 *The Hartree-Fock Method*

The problem of describing at best the ground state of a N-electron problem in terms of a single Slater determinant belongs to the general framework of variational problems, see Appendix B.5. The simple idea is that the average energy $E^{\text{var}}[a] = \langle a|H_{\text{tot}}|a\rangle$ of any state $|a\rangle$ is greater than or equal to that of the ground state. The lower $E^{\text{var}}[a]$ is, the closer $|a\rangle$ gets to the ground state. When for $|a\rangle$ we take a generic Slater determinant, the "best" state in its class is the result of the minimization of the energy

---

[8] This means that the fixed-spin states include a degree of geometric correlation of the electronic motion induced by the symmetry properties of the spatial wavefunction. Employing a basis where the perturbation V_{ee} is diagonal within each unperturbed degenerate space follows the same strategy as the adoption (Appendix B.8.1 and Sect. 2.1.7) of the $|l,\; s,\; j,\; m_j\rangle$ basis to have $H_{\text{s-o}}$ diagonal within the degenerate multiplets.

$$\begin{aligned}
&E^{\text{var}}[\psi_{\alpha_1}, \ldots, \psi_{\alpha_N}] \\
&\quad = {}^A\langle \alpha_1, \ldots, \alpha_N | H_{\text{tot}} | \alpha_1, \ldots, \alpha_N \rangle^A \\
&\quad = \int dw_1, \ldots, dw_N \, \Psi^*_{\alpha_1,\ldots,\alpha_N}(w_1, \ldots, w_N) H_{\text{tot}} \Psi_{\alpha_1,\ldots,\alpha_N}(w_1, \ldots, w_N) \\
&\quad = \sum_i \langle \alpha_i | H_1 | \alpha_i \rangle + \frac{1}{2} \sum_{i,j} \Big[\int dw\, dw' \, |\psi_{\alpha_i}(w)|^2 \, v_{ee}(w, w') \, |\psi_{\alpha_j}(w')|^2 \\
&\qquad - \int dw\, dw' \, \psi^*_{\alpha_i}(w) \psi^*_{\alpha_j}(w') \, v_{ee}(w, w') \, \psi_{\alpha_j}(w) \psi_{\alpha_i}(w') \Big] \qquad (2.74)
\end{aligned}$$

under arbitrary variations of the N single-particle wavefunctions ψ_{α_i} composing the Slater determinant. We only require the ψ_{α_i} to remain mutually orthonormal $\int dw \, \psi^*_{\alpha_i}(w) \psi_{\alpha_j}(w) = \delta_{ij}$. In Eq. (2.74), the "one-particle term" $H_1(w) = \left[-\frac{\hbar^2}{2m_e} \nabla^2_{\mathbf{r}} - \frac{Ze^2}{|\mathbf{r}|} \right] \otimes \mathbb{1}_{\text{spin}}$ would describe the individual motion of each electron in the field of the nucleus, and $v_{ee}(w, w') = \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \otimes \mathbb{1}_{\text{spin}}$ represents the electron–electron Coulomb repulsion of electrons at locations $\mathbf{r}$ and $\mathbf{r}'$.

Finding a minimum of E^{var} is the problem of minimizing a *functional*, i.e. a function whose independent variables are (a set of) functions. This constrained minimization problem is formally solved if the ψ_{α_i} satisfy the set of coupled nonlinear integro-differential equations called *Hartree-Fock* (*HF*) *equations*:

$$\begin{aligned}
&\overbrace{H_1(w)\psi_\alpha(w)}^{1} + \overbrace{\int \sum_\beta |\psi_\beta(w')|^2 \, v_{ee}(w, w') dw' \, \psi_\alpha(w)}^{2} \\
&\quad - \overbrace{\int \sum_\beta \psi^*_\beta(w') \, v_{ee}(w, w') \, \psi_\beta(w) \, \psi_\alpha(w') dw'}^{3} = \varepsilon_\alpha \psi_\alpha(w). \qquad (2.75)
\end{aligned}$$

Each numbered term in Eq. (2.75) derives from a corresponding term in the total energy (2.74). If one pretends that all ψ_β functions are given *known* functions (rather than the unknown functions they really are), then equations (2.74) become *linear* in ψ_α. These equations for ψ_α have then a Schrödinger-like form. Term 1 contains the kinetic energy plus the Coulomb attraction of the nucleus; term 2 represents the Coulomb repulsion of the average charge distribution of all electrons ($\sum_\beta |\psi_\beta(w')|^2$ represents the number-density distribution of the N electrons at position&spin w'); term 3 is a nonclassical nonlocal exchange term which, in particular, removes the unphysical repulsion of the electron with itself introduced by term 2 (observe that the $\alpha = \beta$ terms in the sums of terms 2 and 3 cancel).[9] The HF equations realize a

---

[9] In the β sum of term 3, only $m_{s\,\beta} = m_{s\,\alpha}$ terms survive, as v_{ee} is purely orbital and does not modify spin.

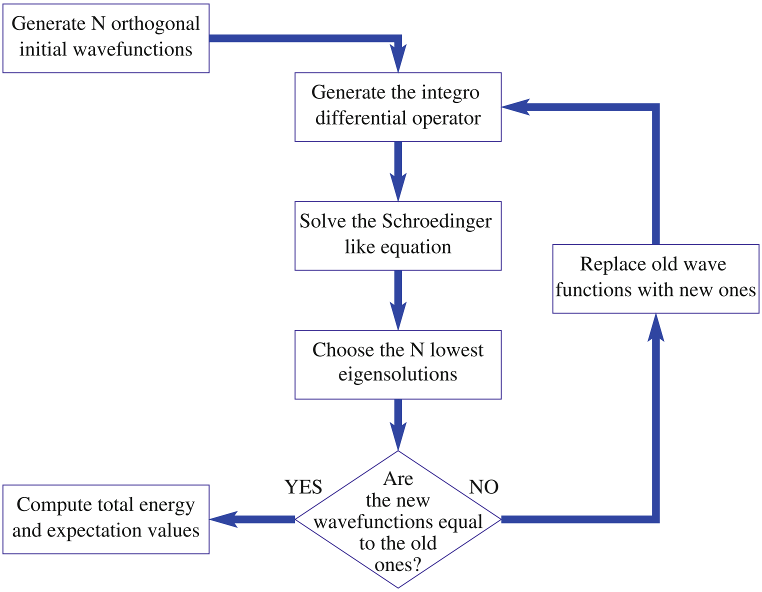


Fig. 2.18 An idealized resolution scheme of the self-consistent HF equations (2.75)

natural way to deal with the electron–electron repulsion as accurately as possible at the mean-field level.

Terms 2 and 3 of Eq. (2.75) depend explicitly on the (unknown) wavefunctions ψ_β. The standard strategy (Fig. 2.18) for the solution of the HF equation is based on initially assuming that all ψ_β in Eq. (2.75) are known: start from some arbitrary initial set of N orthonormal one-electron wavefunctions, put them in place of all ψ_β's in Eq. (2.75), thus generating a first approximation for the effective potential energy acting on the single electrons; solve (usually numerically) the linear equations for ψ_α; from the list of solutions, take the N eigenfunctions with lowest *single-particle eigenenergy* ε_β; re-insert them into the equations (2.75) in place of the ψ_β's thus generating a better approximation for the effective potential energy; iterate this procedure as long as needed. Usually, after several iterations (of the order of 10, depending on the starting ψ_β), *self-consistency* is reached, i.e. the wavefunctions do not change appreciably from one iteration to the next. The converged wavefunctions allow one to compute several observable quantities, and in particular the total HF energy given by Eq. (2.74). The sum of the nuclear potential plus the repulsion of the charge distribution of the other electrons represents the *self-consistent potential* energy V_{HF} driving the motion of the electrons.

Until now, no assumption has been made about the symmetry of $V_{\mathrm{HF}}(w)$, which is safe in a general context. In practice, a simplifying approximation is usually made when applying HF to atoms: that the electron charge distribution and therefore the self-consistent V_{HF} are spherically symmetric functions, like the attraction to the nucleus V_{ne}. This approximation allows one to separate variables in Eq. (2.75), like in the Schrödinger problem for the one-electron atom, and write each HF single-particle solution ψ_α as the product of a radial wavefunction $R_{n\,l}(r)$ times a spherical harmonic $Y_{l\,m_l}$ times a spinor $\chi_{m_s}(\sigma)$, as in Eq. (2.27). In the spherical approximation, one-electron wavefunctions are labeled by hydrogen-like quantum numbers $\alpha = (n, l, m_l, m_s)$: here quantum numbers l, m_l, m_s label exactly the same angular and spin dependence as in Eq. (2.27) for the one-electron atom. In contrast, the radial wavefunctions $R_{n\,l}(r)$ differ from those of Eq. (2.13), and are usually determined numerically. Despite their differences, the HF radial functions and those of the one-electron atom share the following properties: (i) the number of radial nodes $(n-l-1)$ defines n, and (ii) near the nucleus $R_{n\,l}(r) \propto r^l$.

The N-electron ground state is built by filling the single-electron levels starting from 1s, 2s, 2p,… upward. As expected, the spherically symmetric self-consistent potential $V_{\text{HF}}(r)$ felt by each electron behaves as $\simeq (N-1-Z)\,e^2/r$ for large r and as $\simeq -Ze^2/r$ for small r, see Fig. 2.19. Because the potential has not a simple Coulomb shape, the single-electron levels do *not* coincide with the energies of the one-electron atom, Eq. (2.10), and importantly their energy depends on l, not only on n. Indeed, an ns orbital, with larger probability than np near the nucleus (where the effective HF potential is more strongly attractive), is placed lower in energy. Thus the Hartree-Fock method accounts quite naturally for the observed *l-ordering* ns, np, nd,… of the single-electron levels observed in the atomic spectra (e.g. for He in Fig. 2.17). Moreover, the faster-than-Coulombic raise of the effective potential of many-electron atoms induces a n-dependence of the shell energy which is more rapid than the n^{-2} of the one-electron atom.

Figure 2.20 reports the filled single-electron radial distribution associated to the HF wavefunctions for the Ar atom. The typical radii of the individual shells vary

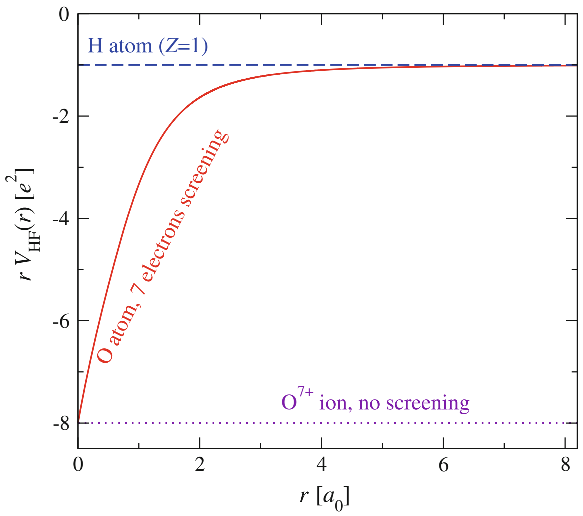


Fig. 2.19 A sketch of the radial dependence of the one-electron Hartree-Fock effective potential energy multiplied by r for atomic oxygen, $N = Z = 8$ (*solid line*). This quantity interpolates between the hydrogen -1 value (*dashed*) at large distance, and the value -8 generated by the bare oxygen nucleus in the context of a O^{7+} ion (*dotted*) at small distance

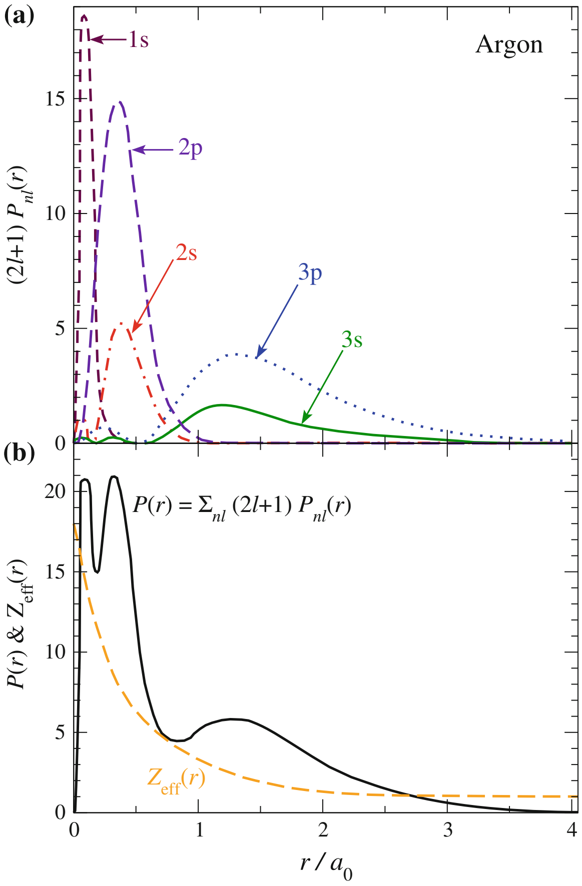


Fig. 2.20 The radial distributions of a many-electron atom (Ar, $N = Z = 18$) computed by means of the HF self-consistent method. **a** The radial probability distribution for each one of the filled single-electron states. Note that the characteristic radius of the innermost shell $n = 1$ is $\approx a_0/Z$, while the outer filled shell ($n = 3$) is slightly larger than a_0. **b** The total radial probability distribution $P(r)$ and effective integrated charge $Z_{\mathrm{eff}}(r)$ generating the effective one-electron HF potential acting on each electron (Data from Ref. [1])

substantially with n, from $\approx a_0/Z$ of 1s, to a few times a_0 for the 3p valence shell. As a result, 3 peaks associated to each of the filled shells emerge prominently in the total probability distribution $P(r)$, Fig. 2.20b.

The independent-electron self-consistent spherical-field HF model has become far more than just an approximation to the actual atomic state: it provides the ordinary language of atomic physics. The electron occupancies of the single-electron orbitals composing the Slater determinant with the largest overlap to the actual correlated many-electron atomic eigenstate are adopted routinely as a label of that eigenstate. For example, the standard notation for the electronic ground-state *configuration* of Mg is $1s^2 2s^2 2p^6 3s^2$.

2.2.5 Electronic Structure Across the Periodic Table

Let us now examine electronic configurations across the periodic table in the light of the HF theory. The $1s^2$ ground configuration of He has total energy $E_{1s^2}^{\text{tot}} = -77.8\,\text{eV}$ according to HF, $\sim 1.2\,\text{eV}$ above the measured ground-state energy; the radial dependence of the one-electron wavefunction is of course nonhydrogenic. In Lithium, a third electron adds into 2s (configuration $1s^2 2s$). The HF ground-state total energy is $-7.4328\,E_{\text{Ha}}$ [5], to be compared with the experimental ($1^{\text{st}} + 2^{\text{nd}} + 3^{\text{rd}}$) ionization energy $7.4755\,E_{\text{Ha}}$, with an error of about 1 eV. The first ionization potential can be computed by subtracting the total energy given by a self-consistent calculation for the positive ion, $N = Z - 1$: for Li, one obtains a ionization potential of 5.34 eV, in good accord with the experimental value 5.39 eV. This value is much smaller than that of He (24.59 eV). The reason is that the binding of the $n = 2$ shell is much weaker than 1s. Beryllium has a $1s^2 2s^2$ ground state. For all these atoms ($N \le 4$) involving only s orbitals, the spherical approximation is appropriate.

Starting from Boron, electrons occupy progressively a degenerate p sub-shell: as the p orbitals are non spherically symmetric, the spherical approximation for the self-consistent field is questionable. The 2p sub-shell is completely filled as Neon ($N = Z = 10$), the next noble gas, is reached. Again Ne is a spherically symmetrical atom, since $\sum_{m_l} |Y_{l\,m_l}(\theta, \varphi)|^2$ is independent of θ and φ. The ionization potential of Ne is again very large, but not as much as that of He (see Fig. 2.21). The next atom, Na, involves one electron in the 3s sub-shell, which is located much higher in energy than 2p. Again the ionization potential has a dip, as shown in Fig. 2.14, which can be interpreted as the starting of a new shell which is only weakly bound. As $Z = N$ further increases, the filling of the $n = 3$ shell proceeds fairly smoothly, with 3s and 3p becoming more and more strongly bound until the next noble gas Ar is reached.[10] For Z this large, the l-dependence of the single-particle HF energy is so strong that the HF self-consistent field puts the 4s sub-shell lower than 3d. Indeed experiment shows that potassium has ground state $1s^2 2s^2 2p^6 3s^2 3p^6 4s$ rather than $1s^2 2s^2 2p^6 3s^2 3p^6 3d$. The physical properties of this atom are similar to those of other alkali metals (Li, Na).

After filling 4s at $Z = 20$ (Ca), and before involving 4p, electrons start to fill the 3d shell. Note however a few inversions (as in Cr and Cu) indicating that 4s and 3d are very close in energy, and subtler effects of electron correlation play a relevant role. Further intersections associated to a strong l-dependence of energy occur as 4d, 4f and 5d are being filled, as reported in the periodic table. Similar properties of all elements with a given number of s or p electrons in the outermost shell suggest the overall arrangement of the periodic table. The "low-energy" properties of atoms

---

[10] For argon, $Z = 18$, the HF approximation finds a total energy of $-526.817\,E_{\text{Ha}}$, which is $0.791\,E_{\text{Ha}} = 21.5\,\text{eV}$ in excess of the experimental energy [21]. The absolute error is rather large, which indicates that the neglect of dynamical correlations in the electronic motion is a serious drawback of HF. However, the relative error is $\sim 0.15\%$ only, and the excess energy per electron amounts to approximately 1 eV, indicating that this mean-field approximation captures the bulk of the electron–electron repulsion.

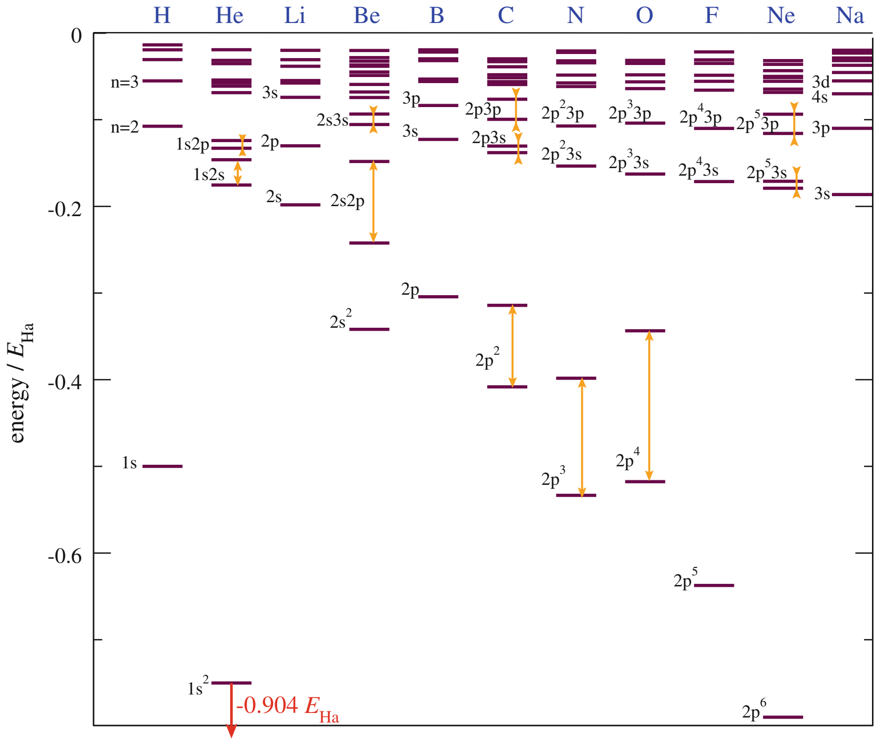


Fig. 2.21 An energy diagram comparing the observed ground state and lowest excitations of the $1 \leq Z \leq 11$ atoms. The zero of the energy scale (*top* of the figure) is set for each atom at the ground energy of the monocation. The energy spread of the atomic multiplets for fixed configuration, highlighted by the *vertical arrows*, is due to Coulomb exchange and correlation energy

with incomplete d shells (transition metals) and f shells (lanthanides) are relatively similar in each group.

The size of the atoms (a not especially well-defined property) computed with HF is in relatively good accord with the empirical trends of Fig. 2.15. In particular, noble gases are especially small and alkali atoms especially large relative to other atoms of similar Z; on the whole, the size of neutral atoms tends to increase slowly with Z. Cations (=positive ions) can be produced with any charge $Z - N$: the size of these species decreases as shells empty up and screening is less and less effective. A cation with N electrons often exhibits the same formal ground electronic configuration as the neutral atom ($Z = N$): this holds especially for small charging, $N \simeq Z$. However all single-electron wavefunctions shrink closer to the nucleus in cations than in the neutral atom. Anions (=negative ions) are stable in gas phase only for certain atoms and only up to a maximum charging of 1 electron (i.e. $N \leq Z + 1$). The HF model signals that a certain ionic configuration is unstable by never reaching self consistence. The HF stability or lack thereof often agrees with experiment.

For example all halogens (F, Cl, Br, I, At), thanks to their almost complete relatively "deep" in energy $n\text{p}^5$ shell (Fig. 2.21), have positive *electron affinity* (defined as the ionization potential of the negative ion), which means that their negative ion is stable against the loss of the extra electron.

Beside ionic states, HF permits also to compute (to some extent) excited states and excitation energies. For example, after computing the ground-state properties of Na, by filling the $N = 11$ lowest single-electron levels as $1\text{s}^2 2\text{s}^2 2\text{p}^6 3\text{s}$, one could run a new self-consistent calculation putting the 11th *valence* electron in 3p rather than in 3s (configuration $1\text{s}^2 2\text{s}^2 2\text{p}^6 3\text{p}$). The self-consistent field turns out different, and the total energy is larger. The difference in total energy between the two calculations is a fair estimate of the excitation energy, here of the 3s→3p transition of Na, approximately 2 eV, see Fig. 2.24.

2.2.6 *Fundamentals of Spectroscopy*

The main observation needed to understand the spectra of many-electron atoms is that the dipole operator driving electromagnetic transitions is a one-electron operator. In concrete, it the sum of the individual dipoles of the single electrons:

$$\mathbf{d} = \sum_i \mathbf{d}_i = -q_e \sum_i \mathbf{r}_i . \tag{2.76}$$

Consider initial and final states described in the HF model as two antisymmetrized states as in Eq. (2.61). Between two such states, the matrix element of one such operator is[11]

$$\begin{aligned}
\langle \beta_1, \ldots, \beta_N|^A \sum_i \mathbf{d}_i |\alpha_1, \ldots, \alpha_N\rangle^A \\
&= \frac{1}{N!} \sum_i \sum_{P\,P'} \langle \beta_{P_1}|\alpha_{P'_1}\rangle \ldots \langle \beta_{P_i}|\mathbf{d}_i|\alpha_{P'_i}\rangle \ldots \langle \beta_{P_N}|\alpha_{P'_N}\rangle \\
&= \frac{1}{N!} \sum_i \sum_{P\,P'} \delta_{P_1,P'_1} \ldots \delta_{P_{i-1},P'_{i-1}} \langle \beta_{P_i}|\mathbf{d}_i|\alpha_{P'_i}\rangle \ldots \delta_{P_N,P'_N} \\
&= \frac{1}{N!} \sum_{i\,P} \langle \beta_{P_i}|\mathbf{d}_i|\alpha_{P_i}\rangle = \sum_i \langle \beta_i|\mathbf{d}_i|\alpha_i\rangle .
\end{aligned}$$

---

[11] The first simplification comes because all one-electron angular products vanish unless all $\beta_{P_k} = \alpha_{P'_k}$. This statement is rigorous under the simplifying assumption that the single-electron basis states composing the initial state $|\alpha_1, \ldots \alpha_N\rangle^A$ are essentially equal to those composing the final state $|\beta_1, \ldots \beta_N\rangle^A$. The second simplification is due to the $N - 1$ Kronecker deltas forcing the P' permutation to match P. The final elimination of the sum over P is due to this sum leading to $N!$ copies of the same matrix elements.

This calculation relates the N-electron electric-dipole matrix element to that of one-electron states. When $\langle \beta_i | \mathbf{d}_i | \alpha_i \rangle$ vanishes, because it violates the one-electron electric-dipole selection rules of Sect. 2.1.9, the overall matrix element vanishes.

In concrete: dipole-allowed transitions occur only between N-electron states differing in *one electron* that *makes a dipole-allowed transition*, with the other $N-1$ electrons remaining in their initial single-particle state. With this condition, any of the N electrons in the initial states $|\alpha_i\rangle$ can make its own transition to any initially empty state. The angular part of all one-electron wavefunction is a standard spherical harmonic Y_{lm}. Therefore the electric-dipole selection rules ($\Delta l = \pm 1$, $\Delta s = 0$) derived for the one-electron atom in Sect. 2.1.9 continue to hold for the electron executing the transition. Here we report a few examples of *allowed* transitions of beryllium: $1s^2 2s^2 \rightarrow 1s^2 2s2p$, $1s^2 2s2p \rightarrow 1s^2 2s4d$, and $1s^2 2s^2 \rightarrow 1s2s^2 3p$; and a few examples of *forbidden* transitions: $1s^2 2s^2 \not\rightarrow 1s^2 2s3d$, $1s^2 2s^2 \not\rightarrow 1s^2 2p^2$, and $1s^2 2s^2 \not\rightarrow 1s2s2p3p$. Further rules discussed in Sect. 2.2.9 restrict dipole-allowed transitions depending on changes of total angular momenta obtained by coupling the spins and orbital angular momenta of individual electrons following the schemes of Sect. 2.2.8.3.

2.2.7 Core Levels and Spectra

A result of HF theory is that the screening of the deep (strongly bound) single-particle levels is scarce, so that their energy is highly negative, and decreases essentially as $\propto -Z^2$. Indeed, in the independent-electron language, it should be possible to excite electrons from the deep *core* levels 1s, 2s,... For example for Na, a configuration such as $1s^{\mathbf{1}} 2s^2 2p^6 3s^{\mathbf{2}}$ could be investigated, where an inner 1s electron has been promoted to the outer 3s shell. Producing this state requires a huge excitation energy (of the order of 40 E_{Ha}), to the extent that one might suspect that such a highly unbound state has no right to exist. Indeed, the atom in this state has plenty of electric-dipole-allowed transitions to get rid of big chunks of this excitation energy, to states at much lower energy, such as $1s^2 2s^2 2p^5 3s^2$. According to Eq. (2.45), the decay transition rate is very large as it grows with the third power of the energy associated to the transition, which dominates over the reduction in dipole matrix element due to the small size of the initial shell, to a total $\sim Z^{4-5}$ dependence—see Eq. (2.55). Accordingly, the spectral broadening due to the short lifetime of *core states* is often huge, exceeding $\hbar\gamma \approx 1$ eV. Despite such huge broadening, core-hole states are not just a theoretical prediction of the independent-electron model, but they are routinely observed in UV and X-ray spectroscopies.

Many experiments probing core spectra with photons can be classified as absorption or emission, with the same conceptual scheme of Figs. 1.3 and 1.4. Absorption data (Fig. 2.22) show a remarkable regularity of the spectra above $\approx$100 eV, and systematic changes of peak positions and intensities as Z is increased. A characteristic feature of X-ray absorption spectra is the asymmetry of the peaks, which exhibit a sharp *edge* at the low-energy side and a broad slow decrease at the high-energy side.

This edge occurs because below the minimum excitation energy for the core state no absorption takes place, while above threshold, the core electron may be promoted to several empty bound and unbound states of the atom (whether in gas or condensed phase), leading to continuum absorption. The slow intensity decay above edge is due to an increase of the ejected-electron kinetic energy (which equals the difference between the absorbed photon energy and the energy of the atomic excited core state): the final state becomes increasingly orthogonal to the initial core level, and correspondingly the dipole matrix element (2.46) decreases. For this same reason, an X-ray photon hitting an atom is much more likely to extract a core electron than a weakly-bound outer-shell electron.

Emission spectra show the same simplicity and regularity as absorption spectra. Initial atomic excitations are usually prepared by collisions with high-energy electrons. The subsequent emission involves transitions only from levels for which enough energy is made available by excitation. For example, if 2 keV electrons are used to excite the sample, emission involving the 1s shell is observed for all $Z \leq 14$ (Si), but not for P and higher-Z atoms (see Fig. 2.22).

Yet another uncalled-for traditional notation haunts core states and X-ray spectra: a *hole* (=missing electron) in shell $n = 1, 2, 3, 4, \ldots$ is labeled $K, L, M, N, \ldots$ The substructures related to states of different l and j acquire a Roman counting subscript (e.g. L_{III} for $2p\,^2P_{3/2}$), as in Fig. 2.23. Dipole-allowed transitions in emission are organized in series according to the initial shell, with a Greek-letter subscript for the final shell. For example, the transition $K \to L$ (in other words the decay $1s2s^22p^6\cdots \to 1s^22s^22p^5\cdots$) produces the K_α emission line, $K \to M$ is K_β, and $L \to M$ is L_α.

In the days of the great discoveries of chemistry and physics, when the structure and classification of atoms were being understood, H. G.-J. Moseley acquired and compared characteristic emission spectra of many elements: he showed that the K_α

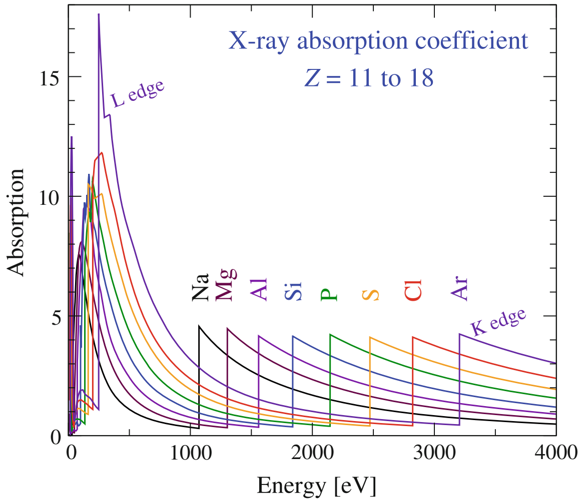


Fig. 2.22 The observed absorption coefficient of all atoms in the third row of the periodic table, showing, for increasing Z, the regular displacement of the K edge (at the *right*), and a gradual buildup of the L edge (at the *left*)

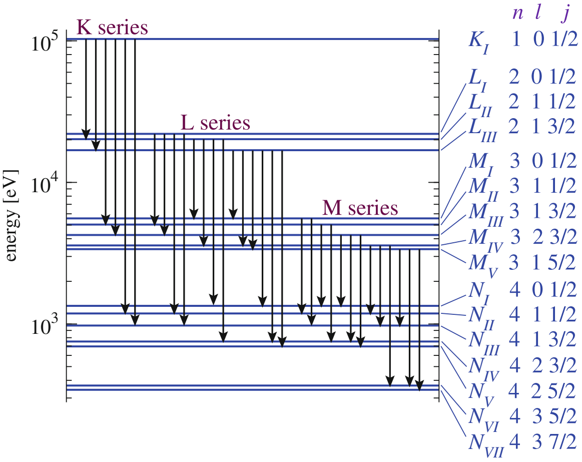


Fig. 2.23 The observed core-level structure of uranium: states are labeled by the quantum numbers n l and j of the core hole; *arrows* highlight the electric-dipole—allowed emission lines. Within each shell, observe the huge l-related and spin-orbit splittings (Inspired by Fig. 9.17 of Ref. [1])

inverse wavelength (equivalently, energy) is roughly proportional to Z^2. This inverse wavelength fits an approximate phenomenological law:

$$\frac{1}{\lambda_{K_\alpha}} \approx C\,(Z-a)^2. \tag{2.77}$$

In this phenomenological dependence, the proportionality constant C is close to $E_{\mathrm{Ha}}/(4\pi\hbar c) \simeq 1.097 \times 10^7\,\mathrm{m}^{-1}$, and the quantity a, accounting for screening, is approximately 2. The discovery of this regularity permitted Moseley to identify the correct value of Z of each atomic species, thus correcting several mistakes in the early attempts to construct the periodic table of the elements.

The decent accuracy of Moseley's fit suggests that one could estimate the core-level energy positions (within, say, 20 %) without going through a full self-consistent HF calculation. Indeed, the energy of a core level is close to that of an hydrogenic state in an effective $-Z_{\mathrm{eff}}\,e^2/r$ Coulomb potential, where the value of Z_{eff} is the average of the effective potential, like the curve of Fig. 2.19, weighted by the radial distribution $P(r)$ of the single-electron wavefunction of that level. Accordingly, the energy of a core shell can be estimated by means of Eq. (2.10), replacing Z with a $Z_{\mathrm{eff}} \simeq Z - 2$ for the K shell, $Z_{\mathrm{eff}} \simeq Z - 10$ for the L shell, and in general $Z_{\mathrm{eff}} \simeq Z$—(number of electrons in inner shells up to and including the target shell).

Nowadays, X-ray spectroscopies are used routinely in research and applications, including position-sensitive analytic tools, local probes of the near chemical environment of different atomic species, and many others. Many more applications of X-ray spectroscopies are reported in the scientific literature and at the web sites of X-ray photon facilities [22].

2.2.8 Optical Spectra

The systematic regularity of the core spectra is lost rapidly when the visible and soft-UV range is explored. Different atoms exhibit vastly different spectra. A few trends and similarities can be recognized.

In the LS scheme, atomic states are labeled $^{2S+1}[L]_J$ to collect their *total angular momenta*. An optional "o" apex can be added to indicate that the overall parity $(-1)^{\sum_j l_j}$ of that state is odd.

2.2.8.1 Alkali Atoms

The alkali atoms have $Z - 1$ electrons filling a number of *close shells* (=completely filled shells) which are energetically well separated (∼20 eV) from the outermost half-occupied n_0s level. The latter is very shallow, with a binding energy ranging from 5.4 eV in Li to 3.9 eV in Cs. The electrons in the inner shells form an essentially "frozen" spherically symmetric core (null spin and orbital angular momentum), which provides an effective potential (Fig. 2.19) for the motion of the outer electron. Like for hydrogen, the ground-state label is $^2S_{1/2}$, because the angular momenta of the alkali atoms coincide with those of the valence electron.

The outer electron is the protagonist of all excitations in the optical spectra of the alkali elements. These spectra resemble the spectrum of a one-electron atom, the main difference being the sizable energy gaps between states characterized by the same n but different l, see Fig. 2.24. In both emission and absorption, dipole-allowed optical transitions appear as diagonal lines in such level schemes, namely s↔p, p↔d, and d↔f, like in Fig. 2.17. A characteristically bright transition in the visible (Li and Na) or near IR (K, Rb, and Cs) spectrum originates from the ∼2 eV n_0s↔n_0p separation, and finds no similarity in hydrogen. This n_0s→ n_0p transition is especially intense due to the large electric-dipole matrix element involving strongly overlapping and fairly extended wavefunctions.

Spin-orbit affects all non-s states. The natural generalization of Eqs. (2.32) and (2.35) to a generic radial potential yields a microscopic estimate of the relevant ξ_{nl} for a given shell:

$$\xi_{nl} = \frac{\hbar^2}{2\, m_e^2\, c^2} \langle n, l | \frac{1}{r} \frac{dV_{\rm eff}(r)}{dr} | n, l \rangle = Z_{\rm eff\, s-o}^4 \alpha^2 E_{\rm Ha} \frac{1}{n^3 l(l+1)(2l+1)} . \quad (2.78)$$

where $Z_{\rm eff\, s-o}$ is implicitly defined by this equation and provides rough estimates of ξ_{nl}. Due to the strong localization of the mean field $r^{-1}\, dV^{\rm eff}(r)/dr$ near the origin, the effective charge $Z_{\rm eff\, s-o}$ is usually larger than that introduced above for the estimate of the level position: $Z_{\rm eff} < Z_{\rm eff\, s-o} < Z$. The spin-orbit level splittings induces splittings of all optical transitions. Table 2.2 reports the observed spin-orbit splittings of the lowest excited p state of the alkali atoms. These splittings are generally larger than those of H: for example the 3p splitting originates a well-resolved

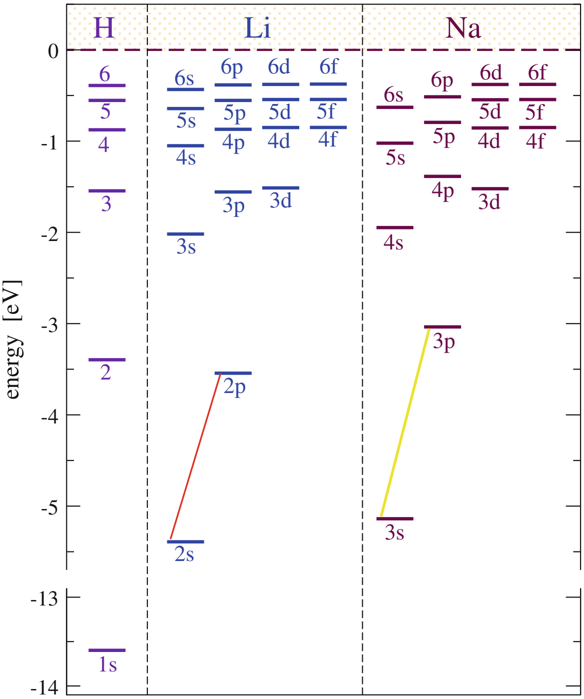


Fig. 2.24 The observed level scheme of lithium and sodium, compared to the levels of hydrogen. For a easier comparison, for each atom the scale reference energy is set at the ground state of its monocation, like in Fig. 2.21. Note the significant dependence of energy not only on n but also on l. The 2s↔2p transition of Li (*red lines* at 670.776 and 670.791 nm) and the 3s↔3p transition of Na (*yellow* "D-lines" at 588.995 and 589.592 nm) are characteristically bright optical transitions in the visible range (Data from Ref. [20])

Table 2.2 The spin-orbit splittings and effective charge of the lowest excited p level of the alkali atoms

Element	H	Li	Na	K	Rb	Cs
Z	1	3	11	19	37	55
Single-electron excited level np	2p	2p	3p	4p	5p	6p
Spin-orbit splitting $\frac{3}{2}\,\xi_{n\mathrm{p}}$ (meV)	0.045	0.042	2.1	7.2	29.5	68.7
$Z_{\mathrm{eff\ s-o}}$	1	0.98	3.5	6.0	10.0	14.2

doublet structure in the optical spectrum of sodium vapors (the characteristic yellow 3s–3p lines of wavelengths 588.99 and 589.59 nm). Remarkably, the spin-orbit splitting of Li 2p is smaller than that of H 2p: the reason is that most of the Li 2p wavefunction lies well outside the compact 1s screening shell.

The fine-structure splittings of higher excited non-s states are smaller than those of the n_0p states, see Fig. 2.25. In the Li spectrum all doublets are split with the smaller J at lower energy, as expected by $\xi_{nl} > 0$, Eq. (2.78). However, the spectra of Na and K display an inverted splitting of the $l \geq 2$ doublets, due to nontrivial exchange and correlation effects (neglected in the mean-field HF approximation) of the valence and core electrons [6] prevailing over the weak spin-orbit interaction.

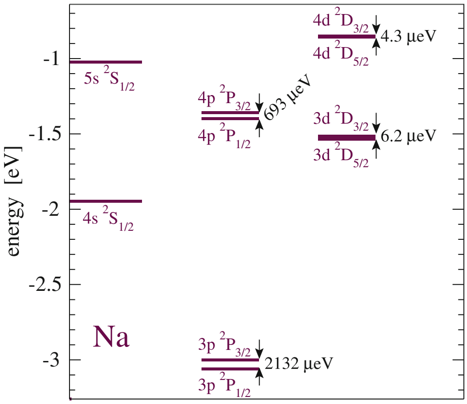


Fig. 2.25 The fine structure of a few low-energy excited levels of sodium. All splittings are widely expanded to make them visible. The regular $P_{1/2}$–$P_{3/2}$ splitting of the p states is governed by the spin-orbit interaction. Note the inverted splitting of the $D_{5/2}$–$D_{3/2}$ doublets. The reference energy is the same as in Fig. 2.24

Within the HF model, Eq. (2.78) can be used to estimate the spin-orbit energy ξ_{nl} for any shell of any atom, not just for the excited shells of alkali atoms. For example, the large effective $Z_{\text{eff}\,\text{s-o}}$ accounts for the colossal spin-orbit splittings (tens or hundreds eV) of the core shells of heavy atoms, observed in X-ray spectra (e.g. the L_{II}–L_{III} splitting shown in Figs. 2.22 and 2.23).

2.2.8.2 Atoms with Elementary Ground States

For a number of atoms, the occupancies of the single orbitals are sufficient to determine uniquely all global symmetry properties of the atomic ground state, in particular its total angular momentum J (thus its degeneracy).

The ground-state symmetry of noble gases, alkali earth and, in general, all atoms in close-shell configurations, including Zn, Cd, Hg, Yb, and No is trivial, as all orbital and spin angular-momentum component cancel: these atoms all qualify as nondegenerate spherically symmetric $^1S_0$.

Likewise, the ground state of alkali metals is simply $^2S_{1/2}$, with a twofold degeneracy associated to spin, but no orbital degeneracy.

B, Sc, and atoms in the same groups IIIB and IIIA, with a single electron in a degenerate p or d shell (and all inner shells complete), are only marginally more complicated: here the total spin and orbital angular momenta equal those of the lone electron. Spin-orbit coupling splits the two levels with $J = L \pm 1/2$, putting $J = L - 1/2$ lower. Accordingly, the ground state of B is $^2P_{1/2}$ and that of Sc is $^2D_{3/2}$.

The last relatively simple class is that of the halogen atoms (group VIIB, p^5 configuration), characterized by a single *hole* in an otherwise full shell. Here, unsurprisingly, this hole carries the same spin and orbital angular momentum ($S = 1/2$ and $L = 1$) as one electron in that shell. However, the effective spin-orbit interaction for

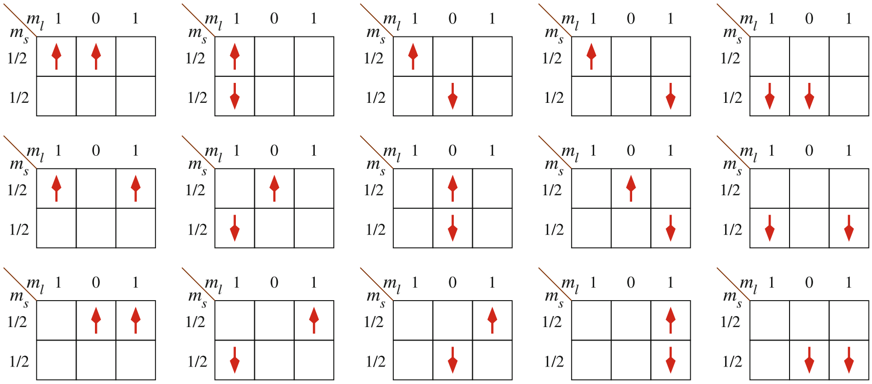


Fig. 2.26 The $6 \times 5/2 = 15$ Slater-determinant states of a p^2 configuration

the hole is reversed in sign.[12] Once this is agreed upon, the ground-state symmetry of all halogens is ${}^2P_{3/2}$. Similarly, the symmetry of the $4f^{13}$ ground state of Tm is ${}^2F_{7/2}$.

2.2.8.3 Atoms with Incomplete Degenerate Shells

Things get intricate when several electrons occupy a degenerate shell, but are too few to fill it completely. The HF independent-electron quantum state usually predicts incorrect degeneracies. For example, the two 2p electrons of carbon can settle in any of the 15 Slater-determinant states sketched in Fig. 2.26. HF puts all the six $M_S = \pm 1$ states of the leftmost and rightmost column at the same energy, and the remaining nine $M_S = 0$ states in the three central columns at a higher energy. This splitting is due to the exchange term in Eq. (2.75). However, the actual ground state of C is *not* 6-fold degenerate! The $M_S = \pm 1$ are the high-$|S_z|$ components of a ${}^3P$ spin and orbital triplet, consisting of $3 \times 3 = 9$ states. The three missing $M_S = 0$ states could be obtained as suitable linear combinations of the nine $M_S = 0$ Slater determinant, which, as such, are not S^2 eigenstates. These combinations are not single Slater determinants, and therefore the HF method cannot be applied to these combinations

---

[12] Observe that for N electrons within a degenerate shell the true spin-orbit operator is $H_{\text{s-o}} = \xi \sum_{i=1}^{N} \mathbf{l}_i \cdot \mathbf{s}_i$. To relate $H_{\text{s-o}}$ to the effective form $H_{\text{s-o eff}} = \xi_{\text{eff}} \mathbf{L} \cdot \mathbf{S}$ that spin-orbit takes in terms of the coupled orbital and spin angular-momentum operators, note that $H_{\text{s-o}} = \xi \sum_{i,j}^{N} \mathbf{l}_i \cdot \mathbf{s}_j - \xi \sum_{i \neq j}^{N} \mathbf{l}_i \cdot \mathbf{s}_j = \xi \left(\sum_{i=1}^{N} \mathbf{l}_i \right) \cdot \left(\sum_{j=1}^{N} \mathbf{s}_i \right) - \xi \sum_{i \neq j}^{N} \mathbf{l}_i \cdot \mathbf{s}_j = \xi \mathbf{L} \cdot \mathbf{S} - \xi \sum_{i \neq j}^{N} \mathbf{l}_i \cdot \mathbf{s}_j$. When the shell is full but for one electron, i.e. there are $N = 2(2l+1) - 1 \equiv d - 1$ electrons, one can rewrite the previous expression as: $H_{\text{s-o}} = \xi \mathbf{L} \cdot \mathbf{S} - \xi \sum_{i \neq j}^{d} \mathbf{l}_i \cdot \mathbf{s}_j + \xi \left(\mathbf{l}_d \cdot \sum_{j}^{N} \mathbf{s}_j + \sum_{i}^{N} \mathbf{l}_i \cdot \mathbf{s}_d \right)$. Observe that the term $\sum_{i \neq j}^{d} \mathbf{l}_i \cdot \mathbf{s}_j = 0$ (since in a completely full shell for a given product $\mathbf{l}_i \cdot \mathbf{s}_j$ there is always another $\mathbf{l}_i \cdot \mathbf{s}_{j'}$ which cancels the previous one) and that $\mathbf{l}_d = -\mathbf{L}$ and $\mathbf{s}_d = -\mathbf{S}$ (since $\mathbf{l}_d = \sum_{i}^{d} \mathbf{l}_i - \sum_{i}^{N} \mathbf{l}_i = 0 - \mathbf{L}$). As a result, $H_{\text{s-o}} = \xi \mathbf{L} \cdot \mathbf{S} + \xi \left[-\mathbf{L} \cdot \mathbf{S} + \mathbf{L} \cdot (-\mathbf{S}) \right] = -\xi \mathbf{L} \cdot \mathbf{S}$. We conclude that the spin-orbit interaction of the "missing" electron, named *hole*, in a halogen atom is the same as that of one electron, but with with reversed spin-orbit $\xi_{\text{eff}} = -\xi < 0$.

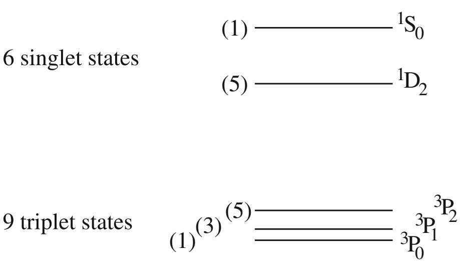


Fig. 2.27 The observed ordering of the p^2 multiplets of carbon and silicon follow Hund's rules. Labels adopt the LS coupling scheme. Degeneracies $(2J + 1)$ are indicated at the *left*

without specific adaptations. The 6 remaining $M_S = 0$ combinations are excited states, identified in Fig. 2.27. The splittings among these states is due to mechanisms discussed in the following.

In general, N electrons can occupy $d = 2(2l+1)$ degenerate spin-orbitals in $\binom{d}{N} = d!/N!/(d-N)!$ different ways, corresponding to physically different (orthonormal) quantum states. The degeneracy of these states is partly (but generally incorrectly) lifted by the HF method. In an actual atom, the splitting of these multiplets is governed by (i) *Coulomb* exchange (mostly accounted for by HF) and correlation (i.e. the effects of the *residual electron–electron interaction* ignored at the HF level, inducing correlated electronic motion), and (ii) the *spin-orbit interaction*.

In the outer atomic shells and for Z smaller than $\sim$30, spin-orbit energy is a comparatively small ($\xi \ll 1\,\text{eV}$) and thus initially negligible interaction. Coulomb repulsion induce much larger splittings of the order of one to several eV (see the He results, Sect. 2.2.3). The Coulomb interaction (and also its residual part, i.e. the part not accounted for by the HF mean field) is spherically symmetric: it thus commutes with total $|\mathbf{S}|^2$ and $|\mathbf{L}|^2$. These operators thus provide a proper labeling of states in the degenerate state. Table 2.3 reports a complete list of the "multiplet" states labeled by their total spin and orbital angular momentum.

Coulomb exchange&correlation acts on the degenerate HF states very much like the full Coulomb repulsion does in He (Sect. 2.2.3). It first of all splits states of different total spin S: low-spin states sit higher in energy because the correlated motion of the electrons takes them, on average, nearer to one another. The ground state will therefore have the highest possible spin: this result agrees with the empirical *first Hund's rule*.

In degenerate configurations there can occur several states of the same S, but different orbital angular momentum L. Electrons avoid one another more efficiently when they rotate all together cooperatively, in states with high L. Accordingly, the state with the largest possible L among those with given S sits lowest in energy. This is known as *second Hund's rule*.

Finally, once the total L and S are determined, the hitherto neglected spin-orbit interaction couples them together to a total angular momentum J. The allowed values of J are given by the usual rule (B.72), and the question of which of them is lowest in energy is decided by the sign of the *effective spin-orbit* parameter for that partly filled shell. While the true spin-orbit parameter is necessarily positive [see Eq. (2.78)],

Table 2.3 All terms into which Coulomb exchange&correlation splits a degenerate $[l]^N$ configuration, for $l = 0$ to 3 and all possible fillings N

$[l]^N$	$S = 0$	1/2	1	3/2	2	5/2	3	7/2
s		^{2}S						
p, p^5		^{2}P						
p^2, p^4	^{1}S D		^{3}P					
p^3		^{2}P D		^{4}S				
d, d^9		^{2}D						
d^2, d^8	^{1}S D G		^{3}P F					
d^3, d^7		^{2}P D F G H		^{4}P F				
		2						
d^4, d^6	^{1}S D F G I		^{3}P D F G H		^{5}D			
	2 2 2		2 2					
d^5		^{2}S P D F G H I		^{4}P D F G		^{6}S		
		3 2 2						
f, f^{13}		^{2}F						
f^2, f^{12}	^{1}S D G I		^{3}P F H					
f^3, f^{11}		^{2}P D F G H I K L		^{4}S D F G I				
		2 2 2 2						
f^4, f^{10}	^{1}S D F G H I K L N		^{3}P D F G H I K L M		^{5}S D F G I			
	2 4 4 2 3 2		3 2 4 3 4 2 2					
f^5, f^9		^{2}P D F G H I K L M N O		^{4}S P D F G H I K L M		^{6}P F H		
		4 5 7 6 7 5 5 3 2		2 3 4 4 3 3 2				
f^6, f^8	^{1}S P D F G H I K L M N Q		^{3}P D F G H I K L M N O		^{5}S P D F G H I K L		^{7}F	
	4 6 4 8 4 7 3 4 2 2		6 5 9 7 9 6 6 3 3		3 2 3 2 2			
f^7		^{2}S P D F G H I K L M N O Q		^{4}S P D F G H I K L M N		^{6}P D F G H I		^{8}S
		2 5 7 10 10 9 9 7 5 4 2		2 2 6 5 7 5 5 3 3				

When several states with the same $^{2S+1}[L]$ occur in the configuration, the number of occurrences is noted below the letter $[L]$ denoting the total orbital angular momentum. In LS (Russell-Saunders) coupling, weak spin-orbit interaction splits further states characterized by different J, e.g. $^2P \rightarrow {}^2P_{1/2}, {}^2P_{3/2}$, not detailed here. After Ref. [5]

the effective parameter for the coupling of total **L** with total **S** may as well be negative, as discussed above for the halogens. Indeed the sign of the effective spin-orbit reverses when more than $2l + 1$ electrons occupy a shell with room for $2(2l + 1)$ electrons. Accordingly, the *third Hund's rule* states that the lowest-energy state has $J = |L - S|$ for a less-than-half-filled shell and $J = L + S$ when the shell is more than half filled.

In the p^2 example, the 15 states of Fig. 2.26 combine themselves into 9 spin-orbit split triplet states $^3P_0$, $^3P_1$, $^3P_2$, plus 6 singlet states $^1D_2$, $^1S_0$. No other state is compatible with Pauli's principle [1]. The Hund-rules ordering of these states is illustrated in Fig. 2.27.

Similar Hund-rules level ordering is usually observed also in configurations involving several incomplete shells, as occur in atomic excited states—see Fig. 2.28b. The observed multiplets of states corresponding to each configuration are scattered

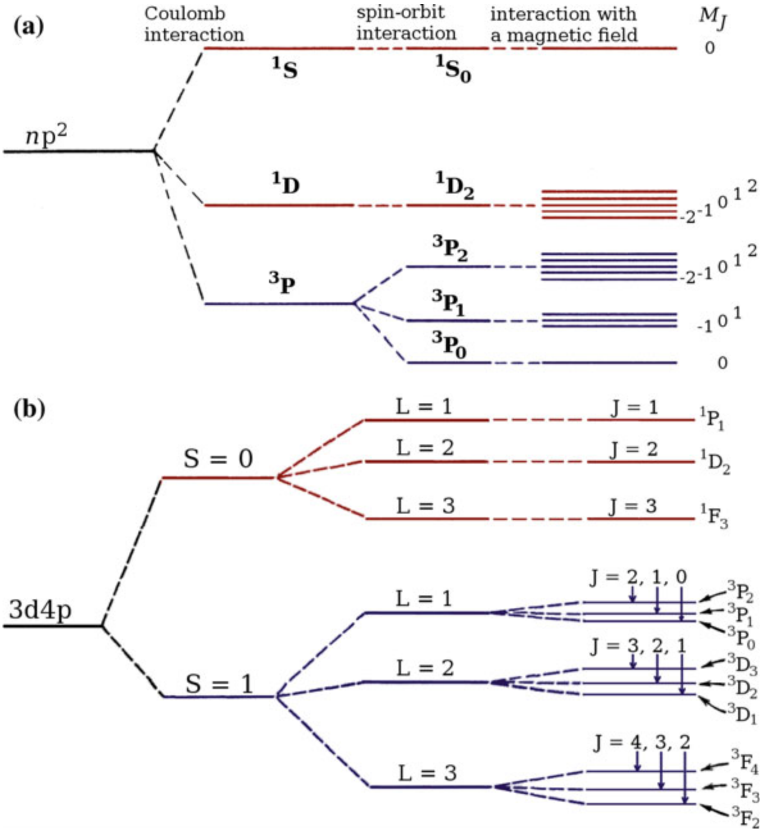


Fig. 2.28 The conceptual sequence of splittings of LS multiplets caused by interactions of decreasing strength. **a** The $6 \times (6-1)/2! = 15$ states of two "equivalent" electrons in a np^2 configuration, including Coulomb and spin-orbit splittings plus those induced by a weak external magnetic field (Zeeman limit). **b** The $6 \times 10 = 60$ states of two "nonequivalent" (different n and/or l) electrons in a 3d4p configuration, as occurs in the spectrum of excited Ti

over ranges shown as brackets in Fig. 2.21 for a few low-Z atoms. The low-energy multiplets scatter out by several eV, while more excited multiplets scatter less, because the Coulomb repulsion is smaller for more extended wavefunctions. Note also that Hund's rules are phenomenological results, not exact laws of nature. Their predictions are occasionally violated. For example, for the $3d^2 4s^2$ configuration of Ti Hund's rules predict an ordering $^3F < {}^3P < {}^1G < {}^1D < {}^1S$, while in reality ^1D lies below ^3P.

The described scheme of coupling all $\mathbf{s}_i$ together to a total $\mathbf{S}$ and all $\mathbf{l}_i$ together to a total $\mathbf{L}$ (followed by spin-orbit coupling of $\mathbf{S}$ and $\mathbf{L}$ together) is called *Russell-Saunders or LS coupling*. It provides a satisfactory basis of coupled states for low-Z atoms, where Coulomb exchange and correlation dominates over spin orbit. For increasing Z the spin-orbit interaction grows rapidly, while electron–electron repulsion remains in the few eV range, and even weakens due to the valence orbitals spreading out. For very large $Z \geq 50$ spin-orbit dominates: $H_{\text{s}-\text{o}}$ must be accounted for before Coulomb terms. The spin-orbit interaction couples the spin and orbital moment of each electron to an individual $j_i = l_i \pm 1/2$. These individual total angular momenta are then coupled to a total J by smaller Coulomb terms. This ordering of the couplings of the angular momenta is called *jj coupling*, and provides another basis for the many-electron states. While the LS basis is almost diagonal for small Z, the jj basis is almost diagonal for large Z (see Fig. 2.29). For intermediate Z, both basis are nondiagonal: the matrix of Coulomb exchange and correlation plus spin-orbit interaction should be diagonalized to express the proper eigenstates as linear combinations of the states of either basis.

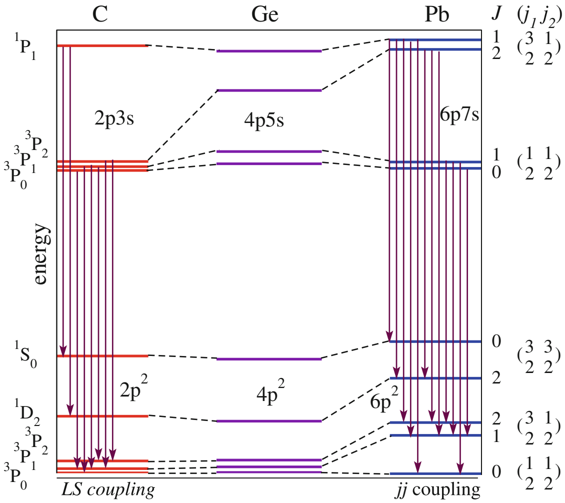


Fig. 2.29 A level correlation diagram, illustrating, for increasing Z, the effect of the increasing relative magnitude of spin-orbit over Coulomb exchange and correlation energies: atomic spectra evolve from the pure LS coupling of carbon ($Z = 6$) to the intermediate coupling (exchange & correlation and spin-orbit of the same order) of germanium ($Z = 32$) to the jj coupling (dominating spin-orbit) of lead ($Z = 82$)

2.2.8.4 Many-Electron Atoms in Magnetic Fields

When a uniform magnetic field acts on a many-electron atom, two very different behaviors are observed depending on whether the atom carries a magnetic moment or not. Atoms with total angular momentum $J = 0$ carry no permanent magnetic dipole available to align with the field: the field induces a small magnetic moment $\sim\mu_B\,(\mu_B B)/\Delta$, where Δ is the energy gap between the ground state and the lowest excitation. We will ignore such tiny effects. Instead, open-shell atoms with $J \neq 0$ carry a magnetic moment $\boldsymbol{\mu} = -g_J \mu_B \mathbf{J}/\hbar$, with a tendency to align to the field.

For LS coupling, the appropriate Landé g-factor g_J is determined by Eq. (B.78), with the total angular momenta J, L, and S in place of the single-electron j, l, and s. As discussed in Sect. 2.1.10, this total magnetic moment, derived by the coupling of orbital and spin contributions, is relevant in the limit of weak external magnetic field (Zeeman limit). In practical experiments this is the relevant limit for many-electron atoms, due to the Z^4 increase of the spin-orbit energy ξ, often exceeding the maximum field accessible in the lab (of the order of 10 T). Also the opposite strong-field (Paschen-Back) limit can be realized, since in highly-excited states, an electron close to dissociation is weakly affected by the nuclear field.

The simplest splitting pattern—three equally spaced lines, corresponding to $\Delta M_J = 1$, 0 and -1—is called regular Zeeman splitting. An example is shown in Fig. 2.30. It occurs when the initial and final g-factors are equal (typically $g = g_L = 1$). This in principle occurs when either $S = 0$ or $L = 0$. However, $L = 0$ is most unlikely, as the electron making the transition changes its l by unity, thus transitions $L_i = 0 \to L_f = 0$ seldom occur. In practice, regular Zeeman splitting is observed in optical transitions between spin-singlet states ($S = 0$). In all other cases, the Zeeman spectrum shows more complicated patterns due to the different initial and final g-factors (anomalous Zeeman spectrum, see Fig. 2.31).

The most direct experiment to investigate the ground-state degeneracy and magnetic moment of many-electron atoms is the Stern-Gerlach one (Sect. 2.1.5). The amount of deflection of the atoms in a field gradient measures the z component of the magnetic moment, and therefore the Landé g-factor g_J, according to Eq. (2.26):

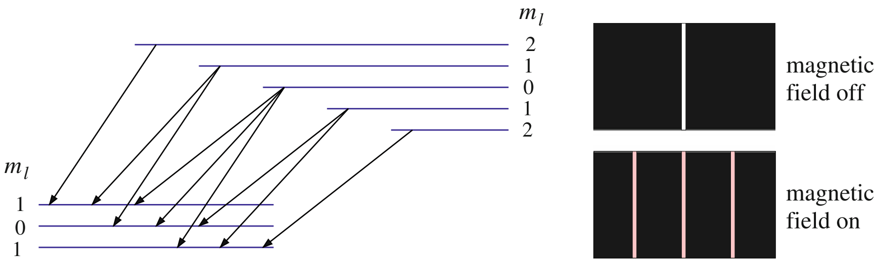


Fig. 2.30 A regular Zeeman spectrum, with its interpretation. It occurs in transitions between $S = 0$ states, which have the same initial and final Landé g-factors. For example, it is observed in the 2s3d $^1D_2 \to$ 2s2p $^1P_1$ emission of Be

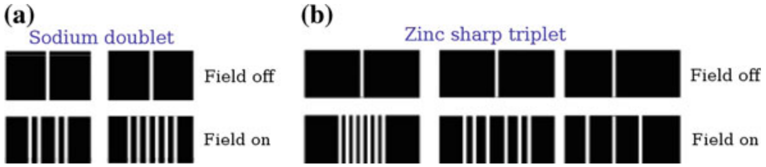


Fig. 2.31 The anomalous Zeeman splitting of Na (**a**) and Zn (**b**) *lines*

$$\mathbf{F}_z = \mu_z \frac{\partial B_z}{\partial z} = -\mu_B \, g_J \, M_J \frac{\partial B_z}{\partial z}. \tag{2.79}$$

The number of sub-beams into which the inhomogeneous field splits the original beam measures directly the number of allowed M_J values, i.e. the ground-state degeneracy $2J+1$.

2.2.9 Electric-Dipole Selection Rules

As discussed in Sect. 2.2.6, the main selection rule requires that a single electron jumps to another state satisfying

$$\Delta l = \pm 1 \qquad \text{(for the one electron making the transition).} \tag{2.80}$$

A few extra electric-dipole selection rules for the total quantum numbers J, L, and S of many-electron atoms in LS coupling also apply. They are summarized below:

$$\text{Parity changes} \tag{2.81}$$

$$\Delta S = 0 \tag{2.82}$$

$$\Delta L = 0, \pm 1 \tag{2.83}$$

$$\Delta J = 0, \pm 1 \qquad \text{(no } 0 \to 0 \text{ transition)} \tag{2.84}$$

$$\Delta M_J = 0, \pm 1 \qquad \text{(no } 0 \to 0 \text{ transition if } \Delta J = 0). \tag{2.85}$$

As both S and L are good quantum numbers only in the limit of very small spin-orbit, in practice selection rules (2.82) and (2.83) are only approximate.

Figure 2.29 draws the allowed transitions in characteristic examples of LS coupling and jj coupling. In this latter scheme, specific dipole selection rules apply, which are described in advanced atomic-physics textbooks [6].

The present Chapter summarizes a few main concepts and experimental evidence in the field of atomic spectroscopy which, in the context of a general course in physics of matter, provide a minimal background and language for understanding and describing the microscopic atomic structure, which lays at the root of the physics of matter. Important conceptual points (e.g. the seniority scheme for the labeling of LS states when L, S and J are not sufficient), and modern spectroscopic techniques (e.g. Auger) are omitted altogether. These and other more advanced subjects, including countless analytic chemical and astrophysical applications of atomic physics, are dealt with in specific textbooks [5–7].

Problems

A ⋆ marks advanced problems.

2.1 A beam of oxygen atoms emerges from an oven with approximately equal populations of its $1s^2 2s^2 2p^2$ triplet states: ${}^3P_0$, ${}^3P_1$, ${}^3P_2$. The average kinetic energy is $E_{\rm kin} = 0.2$ eV. These atoms are sent through a $l = 0.3$ m long Stern-Gerlach magnet, where a field gradient $\partial B/\partial z = 150$ T/m is present. Successively, the beam crosses a region of length $l' = 0.5$ m where the magnetic field is negligible. Compute the total number of beam components detected at the end of the SG apparatus and the distance between the most widely spaced components.

2.2 The two lowest electric-dipole optically allowed transitions starting from the ground state of the Ne atom are observed at energies 16.8 and 19.8 eV. Determine the electronic configuration and the corresponding spectroscopic term of both final excited states involved in these transitions. Describe how these absorption lines change if a uniform magnetic field of 1 T acts on a gaseous Ne sample.

2.3 Construct the scheme of the core levels of Sn, indicating the excitation energies in eV. The K edge is observed at $\lambda_K = 0.425$ Å, and the wavelengths of the first two lines of the K series are: $K_\alpha = 0.517$ Å, $K_\beta = 0.437$ Å. Evaluate the minimum excitation energy needed to observe the L series emission lines after excitation. Evaluate the maximum kinetic energies of photoelectrons produced by the excitation of shells K, L, M of Sn induced by 32 keV X rays.

2.4 Two hydrogen atoms are excited in the quantum states described by the following wavefunctions:

$$(a) \qquad \psi_{300} = \frac{1}{9\sqrt{3\pi}\, a_0^{3/2}} \left(3 - 2\frac{r}{a_0} + \frac{2}{9}\frac{r^2}{a_0^2} \right) e^{-r/3a_0}$$

and

$$(b) \qquad \psi_{210} = \frac{1}{4\sqrt{2\pi}\, a_0^{3/2}} \frac{r}{a_0} e^{-r/2a_0} \cos\theta\,,$$

where a_0 is the Bohr radius. For each, evaluate the decay probability rate γ to the ground state

$$\psi_{100} = \frac{1}{\sqrt{\pi}\, a_0^{3/2}} e^{-r/a_0}$$

in the electric-dipole approximation.

[Recall that in this approximation $\gamma_{if} = (3\pi\varepsilon_0)^{-1}\hbar^{-4}c^{-3}\,(\mathcal{E}_i - \mathcal{E}_f)^3\,|\langle f|\mathbf{d}|i\rangle|^2$, where the Cartesian components of $\mathbf{d}$ are $-q_e r\sin\theta\cos\phi$, $-q_e r\sin\theta\sin\phi$, and $-q_e r\cos\theta$.]

2.5 The 7 valence electrons of the Fe$^+$ ion in its ground state assume the configuration 3d^6 4s. Evaluate the ground-state magnetic moment and the number of components in which the spin-orbit interaction splits the lowest-energy (according to the first and second of Hund's rules) degenerate configuration.

2.6 Determine the number of absorption lines that a gas-phase sample of atomic chlorine at 1,000 K exhibits in the 3s^{2}3p^5(^{2}P) → 3s^{2}3p^{4}4s(^{2}P) transition. Assume a significant thermal population of the excited state of the 3s^{2}3p^5(^{2}P) configuration, given that this excited state sits 109.4 meV above the ground state. If this sample is immersed in a static uniform magnetic field $|\mathbf{B}| = 1.5$ T, how many distinct sub-lines does each of the lines determined above split into?

2.7 The three components of the ground configuration 3d^{2}4s^2 ^{3}F of atomic titanium are at energies 0, 0.02109, and 0.04797 eV. The absorption spectrum of Ti vapor at 1,000 K involves lines related to transitions starting from the ground-state components to 3d^{2}4s4p levels, which are organized as three groups of states of total orbital angular momentum $L = 2$, 3 and 4. Establish the quantum numbers of the final states of the electric-dipole allowed transitions in each of the three groups. Considering now uniquely the transitions starting from the ground state ^{3}F$_2$, identify the most intense transition in each group.

2.8* Evaluate the magnitude of the magnetic moment $|\boldsymbol{\mu}|$ of atomic Sc, V, and Mn in their respective ground states 3d4s^2, 3d^{3}4s^2, 3d^{5}4s^2. Beams of such atoms with equal kinetic energy are sent through one Stern-Gerlach apparatus. Identify for which of these three elements the least-deflected beam component is deflected the most.

2.9 A 30 keV X-ray beam hits a palladium target. Among the photoemitted electrons several are measured at values of kinetic energy: 5,650, 26,396, 26,670, and 26,827 eV. Given these data, evaluate the core-shell energies of Pd, assign the departing shell for the electrons of the smallest kinetic energy and evaluate the effective charge Z_{eff} (with three significant digits) relevant for the motion of an electron in that shell, consistently with the measured photoelectron energies.

2.10 Helium atoms in the excited state $1s2s(^3S_1)$ are generated and accumulated in a storage vessel for (on average) one hour. They leave it through a tiny orifice at an average speed 2,000 m/s, and into a Stern-Gerlach apparatus characterized by a length 0.5 m and a field gradient $\partial B/\partial z = 9\,\mathrm{T/m}$. Given that the lifetime τ of the metastable $1s2s(^3S_1)$ state is 8,000 s, evaluate the fraction of deflected atoms with a $M_J = 1$ component of total angular momentum over the total number of atoms leaving the storage vessel in a given time. Evaluate also the angular deflection of that component.

2.11 A 10 keV photon beam hits a cesium target. Part of the photoemitted electrons is collimated by suitably placed slits and sent into a transverse-field analyzer characterized by $|\mathbf{B}| = 0.1\,\mathrm{T}$. Electrons circulating with radii 2.3 and 3.2 mm. Based on these data, assign the core shells from which the electrons are emitted and evaluate their excitation energy.

2.12 * In the standard formulation of the hydrogen-atom problem, the nucleus is taken as a point charge producing a potential energy $V_{\mathrm{Coul}}(r) = -e^2/r$ for the electron at distance r. In fact the nucleus has a finite size. Modeling it as a uniform sphere of positive charge, of radius $r_n = 0.9\,\mathrm{fm}$, the potential energy changes to $V_{\mathrm{true}}(r) = -3e^2/(2r_n) + e^2r^2/(2r_n^3)$ inside the nucleus (no change outside, of course). Evaluate the effect of this modification on the hydrogen ground-state energy to first order in perturbation theory. [Suggestion: approximate $\exp(x) \simeq 1$ when $|x| \ll 1$.]

Chapter 3
Molecules

Molecules, particularly diatomic ones, are the simplest systems where several nuclei are bound together by their interaction with one or several electrons. Their simplicity makes diatomics the ideal playground to approach two central concepts of condensed-matter physics: the adiabatic separation of the electronic and nuclear motions, and chemical bonding.

3.1 The Adiabatic Separation

The full Hamiltonian (1.1) for a piece of matter and all of its eigenfunctions Ψ depend on the coordinates r of all electrons and R of all nuclei. This entangled dependency makes the resulting information content catastrophically and often needlessly complicated. A disentanglement of the fast electron dynamics from the slow motion of the nuclei is a crucial conceptual step to make progress in understanding and interpreting an immense body of phenomena and observations regarding matter.

Since the electron mass is much smaller than the nuclear mass, the *ionic and electronic motions should occur over different* (and hopefully decoupled) *time scales*. As masses appear at the denominators of the kinetic terms Eqs. (1.2) and (1.3), electrons move much faster than atomic nuclei. The fast electrons should be capable to follow the slow displacements of the nuclei with essentially no delay.

To implement this decoupling consider the following factorization of the total wavefunction:

$$\Psi(r, R) = \Phi(R)\, \psi_e(r, R). \tag{3.1}$$

Assume that the *electronic wavefunction* $\psi_e(r, R)$ is a solution $\psi_e^{(a)}(r, R)$ of the following *electronic equation*:

$$[T_e + V_{ne}(r, R) + V_{ee}(r)]\, \psi_e^{(a)}(r, R) = E_e^{(a)}(R)\, \psi_e^{(a)}(r, R), \tag{3.2}$$

N. Manini, *Introduction to the Physics of Matter*,
Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-14382-8_3

where (a) represents the set of quantum numbers characterizing a given N-electron eigenstate with energy $E_e^{(a)}$. The electronic eigenfunction $\psi_e^{(a)}(r; R)$ describes an electronic eigenstate compatible with a fixed geometrical configuration R of the ions: the R-dependence of $\psi_e^{(a)}(r; R)$ is purely parametric [no ∇_R operators in Eq. (3.2)]. The factorization defined by Eqs. (3.1) and (3.2) (known as the *adiabatic* or Born-Oppenheimer scheme) implies the assumption that, once an initial electronic eigenstate is selected, the atomic nuclei move slowly enough not to induce transitions to other electronic states. These transitions are prevented by the electronic energy gaps usually being much larger than the typical energies associated with the slow motion of the nuclei.

To make use of the *ansatz* (3.1), observe that $T_e \sim \nabla_r^2$ does not act on the R coordinates:

$$T_e\,[\Phi(R)\,\psi_e(r,R)] = \Phi(R)\,[T_e\psi_e(r,R)]. \tag{3.3}$$

Derivation is slightly more intricate for the nuclear kinetic term:

$$\begin{aligned}\nabla_R^2[\Phi(R)\,\psi_e(r,R)] = \psi_e(r,R)\,\nabla_R^2\Phi(R) + 2\,[\nabla_R\psi_e(r,R)]\,\nabla_R\Phi(R)\\ +\ \Phi(R)\,\nabla_R^2\psi_e(r,R).\end{aligned} \tag{3.4}$$

Thus, when we substitute the factorization $\Psi(r,R) = \Phi(R)\,\psi_e(r,R)$ into the Schrödinger equation $H_{\text{tot}}\Psi = E_{\text{tot}}\Psi$ we obtain

$$\begin{aligned}&-\sum_\alpha \frac{\hbar^2}{2M_\alpha}\Big\{2\,[\nabla_{R\alpha}\psi_e(r,R)]\,\nabla_{R\alpha}\Phi(R) + \Phi(R)\nabla_{R\alpha}^2\psi_e(r,R)\\ &\qquad\qquad +\psi_e(r,R)\nabla_{R\alpha}^2\Phi(R)\Big\}\\ &+\ \Phi(R)\,T_e\psi_e(r,R) + [V_{ne}+V_{ee}+V_{nn}]\,\Phi(R)\psi_e(r,R) = E_{\text{tot}}\,\Phi(R)\psi_e(r,R),\end{aligned} \tag{3.5}$$

(omitting the coordinate dependence of the V terms). We can rearrange Eq. (3.5) as follows:

$$-\sum_\alpha \frac{\hbar^2}{2M_\alpha}\Big\{2\,[\nabla_{R\alpha}\psi_e(r,R)]\,\nabla_{R\alpha}\Phi(R) + \Phi(R)\,\nabla_{R\alpha}^2\psi_e(r,R)\Big\} \tag{3.6}$$

$$\begin{aligned}&+\psi_e(r,R)\left[-\sum_\alpha \frac{\hbar^2\nabla_{R\alpha}^2}{2M_\alpha}\right]\Phi(R) + \Phi(R)\,[T_e+V_{ne}+V_{ee}+V_{nn}]\,\psi_e(r,R)\\ &\qquad = E_{\text{tot}}\,\Phi(R)\psi_e(r,R),\end{aligned} \tag{3.7}$$

in order to highlight the electronic Hamiltonian $(T_e + V_{ne} + V_{ee})$ of Eq. (3.2).

The two terms of line (3.6), involving derivatives ∇_R of the electronic wavefunction $\psi_e(r, R)$ are called *nonadiabatic terms*. These nonadiabatic terms (which, recall, are of nuclear kinetic nature) are much smaller (usually by a factor $\approx m_e/M_\alpha$) than the typical differences between electronic eigenenergies of Eq. (3.2), dominating Eq. (3.7). The *adiabatic approximation* consists precisely in the neglect of the terms of line (3.6).

Substitute a solution $\psi_e^{(a)}(r, R)$ of Eq. (3.2) for $\psi_e(r, R)$ in the remaining terms of Eq. (3.7):

$$\psi_e^{(a)}(r, R)\left[-\frac{\hbar^2}{2}\sum_\alpha \frac{\nabla^2_{R\,\alpha}}{M_\alpha} + E_e^{(a)}(R) + V_{nn}(R)\right]\Phi(R)$$
$$= E_{\text{tot}}\,\psi_e^{(a)}(r, R)\,\Phi(R), \qquad (3.8)$$

where the electronic wavefunction $\psi_e^{(a)}(r, R)$ is displaced to the left of the operator, to stress that the differential part only acts on the ionic part $\Phi(R)$. Note that all three terms T_e, V_{ne} and V_{ee} are entirely (and in principle exactly) accounted for by the electronic eigenvalue $E_e^{(a)}(R)$, the inter-nuclear repulsion V_{nn} remains indicated explicitly, and only the nuclear kinetic term T_n is treated approximately, due to the neglect of nonadiabatic corrections of line (3.6). The electronic wavefunction can now be dropped [formally by multiplying Eq. (3.8) by $\psi_e^{(a)\,*}(r, R)$ and integrating over all electronic coordinates r], to derive the equation for the adiabatic motion of the nuclei described by $\Phi(R)$:

$$\left[-\frac{\hbar^2}{2}\sum_\alpha \frac{\nabla^2_{R\,\alpha}}{M_\alpha} + E_e^{(a)}(R) + V_{nn}(R)\right]\Phi(R) = E_{\text{tot}}\,\Phi(R). \qquad (3.9)$$

The **electronic equation** (3.2) describes the motion of all electrons in the piece of matter. When the number of electrons is greater than one (as is usually the case), Eq. (3.2) involves at least the same technical difficulties as many-electron atoms discussed in Sect. 2.2, made worse by the lack of spherical symmetry. Equation (3.2) is often solved within an approximate quantum many-body method, e.g. the Hartree-Fock method sketched in Sect. 2.2.4, or the density-functional theory [23]. In the polyatomic context, theory faces the extra difficulty that the electronic problem depends explicitly on the position R of all nuclei, through V_{ne}. One should thus solve the electronic problem for many geometric arrangements (classical *configurations* R) of the nuclei, to obtain a detailed knowledge of the parametric R-dependence of the electronic eigenfunction $\psi_e^{(a)}(r, R)$ and eigenvalue $E_e^{(a)}(R)$. The latter is especially fundamental, as it drives the motion of the nuclei through Eq. (3.9). Within the adiabatic approximation, once the electronic eigenstate (a) is chosen, this state never mixes with other eigenstates (a'): the electronic state follows adiabatically the slow nuclear motion. In reality, the neglected nonadiabatic terms originate a small probability that the motion of the nuclei induces transitions to different electronic states.

Once the electronic eigenvalue $E_e^{(a)}(R)$ is obtained, the **Schrödinger equation for the nuclei** (3.9) describes the displacements of the nuclei as driven by a *total adiabatic potential energy*

$$V_{\mathrm{ad}}^{(a)}(R) = V_{nn}(R) + E_e^{(a)}(R), \tag{3.10}$$

which is the sum of the direct Coulombic ion-ion repulsion V_{nn} plus the electronic eigenvalue $E_e^{(a)}(R)$ (which changes as a function of the ionic coordinates, as discussed above). This second term, the "adiabatic electronic contribution", acts as a "glue" which keeps the atoms together. $E_e^{(a)}(R)$ is generally a complicated function of all ionic coordinates R: contrary to V_{nn}, $E_e^{(a)}(R)$ can seldom be expressed as a simple sum of two-body contributions.

Of all the electronic eigenstates labeled by (a), the *electronic ground state* generates the especially important *lowest adiabatic potential energy surface* $V_{\mathrm{ad}}^{(\mathrm{gs})}(R)$, in short $V_{\mathrm{ad}}(R)$. Electronic excitations (a) generate different adiabatic potential energy surfaces.

In the adiabatic scheme, the total adiabatic potential $V_{\mathrm{ad}}(R)$ guides the displacements of the nuclei. Since in common language "nuclear" recalls the internal dynamics of nuclei, in practice $V_{\mathrm{ad}}(R)$ is referred to as the potential energy governing the motion of "*atoms*", or "*ions*".

As a function of the $3N_n$ ionic coordinates, $V_{\mathrm{ad}}(R)$ exhibits two general symmetries, consequences of the symmetries of the original Hamiltonian (1.1) for an isolated system:

- *Translational symmetry*: if all ions are displaced by an arbitrary (equal for all) translation $\mathbf{u}$, then V_{ad} remains unchanged. In formula: $V_{\mathrm{ad}}(\mathbf{R}_1+\mathbf{u}, \mathbf{R}_2+\mathbf{u}, \ldots, \mathbf{R}_{N_n}+\mathbf{u}) = V_{\mathrm{ad}}(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_n})$.
- *Rotational symmetry*: if all ions are rotated by an arbitrary (equal for all) rotation A around a given arbitrary point in space, then V_{ad} also remains unchanged. In symbols: $V_{\mathrm{ad}}(A\mathbf{R}_1, A\mathbf{R}_2, \ldots, A\mathbf{R}_{N_n}) = V_{\mathrm{ad}}(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_n})$.

These symmetries indicate that V_{ad} depends on the *relative positions* of the atoms.

$V_{\mathrm{ad}}(R)$ often shows a well-defined minimum, for the atoms placed at specific relative positions R_{M}. In the simplest example of a di-atom, $V_{\mathrm{ad}}(R)$ depends on the distance $R = |\mathbf{R}_1 - \mathbf{R}_2|$ of the two nuclei. For any two atoms picked at random from the periodic table, the potential energy as a function of R shows the qualitative shape of Fig. 3.1, with an equilibrium inter-ionic separation R_{M} at which $V_{\mathrm{ad}}(R)$ is minimum. If the potential well is deep enough, then at low temperature the ionic motion is confined to a neighborhood of the *equilibrium position* R_{M}: the ions execute small oscillations around R_{M}.

The ionic motion is sometimes treated within *classical* mechanics [$M_n \frac{d^2R}{dt^2} = -\nabla_R V_{\mathrm{ad}}(R)$]. This does not mean that the actual ionic motion is any classical, only

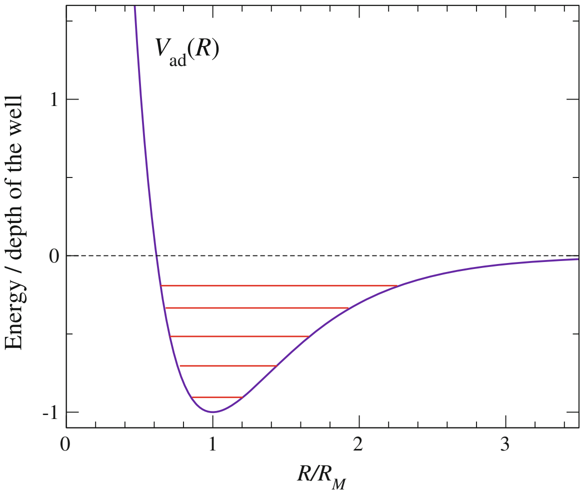


Fig. 3.1 The typical qualitative shape of the adiabatic potential for a diatom, as a function of the inter-ionic separation R. The zero of energy has been taken as the sum of the total energies of the two individual atoms at large distance. *Horizontal lines* represent possible vibrational ground and low-energy excited levels. The actual number and positions of these states depends also on the diatom reduced mass

that under certain conditions (e.g. heavy ions), the classical limit can provide a fair approximation to the actual quantum dynamics. The classical ionic dynamics often yields useful insight into the quantum solution of Eq. (3.9). For example, the classical problem of the normal modes describing *small independent harmonic oscillations* around R_{M} is mapped to that of a set of *quantum* harmonic oscillators (only one oscillator in the case of a diatomic molecule). In Sects. 3.3 and 5.3, we shall return to this fundamental problem in mechanics, and review its solutions.

In summary, the adiabatic scheme provides a fundamental separation of the coupled electron-ion dynamics into two conceptually and practically distinct problems: the electronic equation (3.2) governs the motion of the fast electrons in the field of the ions, imagined as instantaneously frozen; the multidimensional Schrödinger equation (3.9) for the slower motion of the ions is controlled by the adiabatic potential, which is the sum of the ion-ion Coulomb repulsion plus a "gluing" term provided by the electronic total energy—Eq. (3.10).

3.2 Chemical and Non-chemical Bonding

In Fig. 3.1 we anticipated a typical profile for the adiabatic potential of a diatom. If that sort of long-distance attractive and short-distance strongly repulsive behavior is really general, and persists for $N_n > 2$ atoms, then it can provide the microscopic mechanism allowing collections of atoms to bind together, forming all kinds of bound states (including the molecular gases, liquids, and solid objects of everyday experience), with no tendency to collapse to infinitely dense point-like objects. To obtain both qualitative and quantitative insight into the bonding nature of $V_{\text{ad}}(R)$, we investigate conveniently simple model systems.

3.2.1 H_2^+

This investigation starts naturally with the simplest di-atom: H_2^+, which is free of the complications of electron-electron repulsion. We construct a piece of evidence that the adiabatic energy of H_2^+ has a qualitative dependence of the inter-proton distance R of the kind sketched in Fig. 3.1: we verify that, as R decreases from infinity, the lowering in electronic energy $E_e^{(a)}(R)$ wins against the increase in inter-nuclear Coulomb repulsion $V_{nn}(R) = \frac{e^2}{R}$, with a net bonding behavior of the total adiabatic potential energy $V_{\mathrm{ad}}^{(a)}(R)$, Eq. (3.10).

We start considering the potential energy V_{ne} acting on the electron of H_2^+. V_{ne} is the sum of the two attractive Coulomb terms $V_L + V_R$ produced by the *L*eft and *R*ight nuclei respectively, which, in the adiabatic scheme, we take as stationary at a fixed separation R. In the end we will vary R to explore the adiabatic potential-energy function $V_{\mathrm{ad}}(R)$. Figure 3.2 shows three "cuts" of $V_{ne}(r, R)$ as a function of the electron position r. Observe here that in the intermediate region, between the two nuclei, the total potential is roughly twice more negative than it would be if only one of the two nuclei was present. This suggests that the electron moving in the field of both nuclei could take advantage of both attractions and lower its average potential energy by spending a significant fraction of its probability distribution in this intermediate extra-attractive region. We check if this mechanism generates bonding, i.e. if $V_{\mathrm{ad}}(R)$ decreases below its infinite-R value $V_{\mathrm{ad}}(\infty) = E_{1\mathrm{s}} = -E_{\mathrm{Ha}}/2$, namely the energy of one isolated hydrogen atom (no mass correction μ/m_e, because the nuclei are placed at fixed positions!).

To estimate if any energy can be gained, we use a simple variational approach (see Appendix B.5) based on "trial" states build by linear combinations of two states

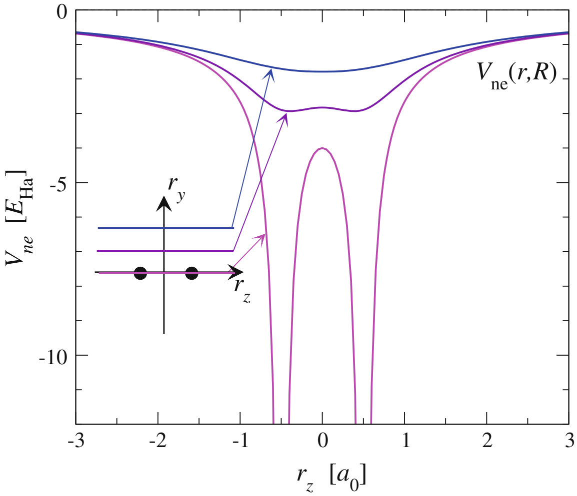


Fig. 3.2 Three parallel cuts of the nuclei-electron attraction $V_{ne}(r, R) = -e^2 [|\mathbf{r} - R\hat{\mathbf{z}}/2|^{-1} + |\mathbf{r} + R\hat{\mathbf{z}}/2|^{-1}]$ in H_2^+, drawn along the *line* through the nuclei (the z axis) and along *parallel lines* at distances $R/2$ and R from the nuclei. Here $R = a_0$

only, namely the two 1s hydrogen orbitals, Eq. (2.21), $|1\mathrm{s}\,L\rangle$ and $|1\mathrm{s}\,R\rangle$ centered at the Left and Right nucleus respectively. Guided by the reflection symmetry of $V_{ne}(r, R)$ across the xy plane, build two normalized trial electronic wavefunctions as $\psi_e^{1\mathrm{s}\,\substack{S\\A}}(r) = \langle r|\substack{S\\A}\rangle$, with

$$
\begin{aligned}
|S\rangle &= \frac{1}{[2(1+\mathrm{Re}\langle 1\mathrm{s}\,L|1\mathrm{s}\,R\rangle)]^{1/2}}\,(|1\mathrm{s}\,L\rangle + |1\mathrm{s}\,R\rangle) \quad \text{Symmetric} \\
|A\rangle &= \frac{1}{[2(1-\mathrm{Re}\langle 1\mathrm{s}\,L|1\mathrm{s}\,R\rangle)]^{1/2}}\,(|1\mathrm{s}\,L\rangle - |1\mathrm{s}\,R\rangle) \quad \text{Antisymmetric.}
\end{aligned}
\tag{3.11}
$$

Figure 3.3 sketches the wavefunctions $\psi_e^{1\mathrm{s}\,S}(r)$, $\psi_e^{1\mathrm{s}\,A}(r)$, with the ingredients $\psi_e^{1\mathrm{s}\,\substack{L\\R}}(r)$.

For any given inter-nuclear separation R, the variational principle guarantees that the ground-state energy $E_e^{(GS)}(R) \leq \langle S|T_e + V_{ne}|S\rangle \equiv \mathscr{E}_S$, and similarly that $E_e^{(GS)}(R) \leq \mathscr{E}_A$. For infinitely large inter-nuclear separation R, $|S\rangle$ and $|A\rangle$ are exact eigenstates, of electronic energy $\mathscr{E}_{\substack{S\\A}}(R=\infty) = E_{1\mathrm{s}} = -1/2\,E_{\mathrm{Ha}}$. For finite R, their mean electronic energy becomes

$$
\mathscr{E}_{\substack{S\\A}} = \langle \substack{S\\A} |T_e + V_{ne}| \substack{S\\A} \rangle = \frac{(\langle L|T_e + V_{ne}|L\rangle \pm \langle R|T_e + V_{ne}|L\rangle) + (L \leftrightarrow R)}{2(1 \pm \langle L|R\rangle)},
\tag{3.12}
$$

where we have omitted the 1s labels and the Re() (the overlap $\langle L|R\rangle$ is real). The $(L \leftrightarrow R)$ terms obtained by exchanging left and right equal the previous ones, thus we can omit them together with the factor 2 at the denominator. Substitute $V_L + V_R$

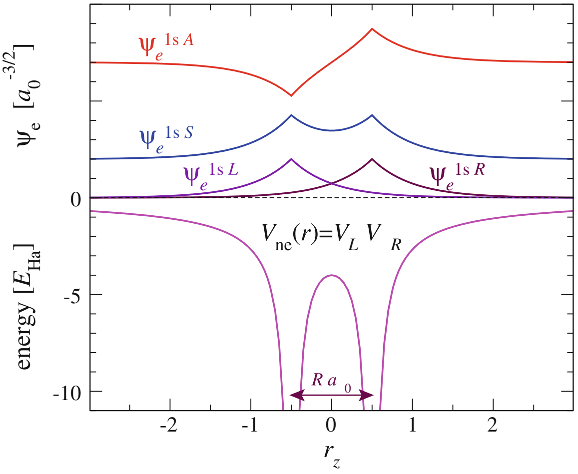


Fig. 3.3 A cut along the molecular axis for: (*lower* to *upper*) the potential energy $V_{ne}(r, R)$, the 1s eigenfunctions of the left $\psi_e^{1\mathrm{s}\,L}(r)$ and the right $\psi_e^{1\mathrm{s}\,R}(r)$ individual potential wells, plus their symmetric $\psi_e^{1\mathrm{s}\,S}(r)$ and antisymmetric $\psi_e^{1\mathrm{s}\,A}(r)$ combinations (shifted upward for better visibility)

for V_{ne}, and reorganize the six matrix elements, to take advantage of the fact that $|L\rangle$ is the ground state of $(T_e + V_L)$:

$$\begin{aligned}\mathscr{E}_{\substack{S\\A}} &= \frac{\langle L|T_e + V_L|L\rangle + \langle L|V_R|L\rangle \pm \langle R|T_e + V_L|L\rangle \pm \langle R|V_R|L\rangle}{1 \pm \langle L|R\rangle} \\ &= -\frac{E_{\text{Ha}}}{2} + \frac{\langle L|V_R|L\rangle \pm \langle R|V_R|L\rangle}{1 \pm \langle L|R\rangle}. \end{aligned} \tag{3.13}$$

The second term contains two Coulomb matrix elements, which are both real and negative and depend on the inter-nuclear distance. Despite the denominator, the symmetric state $|S\rangle$ is lower in energy than $|A\rangle$. The $\langle L|V_R|L\rangle$ term represents the attraction that the right nucleus exerts on the electron sitting around the left nucleus. At large R this attractive term balances almost exactly the $V_{nn} = e^2/R$ inter-nuclear repulsion. This cancellation represents the classical result that the electrostatic interaction energy of a spherically symmetrical neutral object (the H atom) and a remote point charge (the H^+ ion) decays much faster than R^{-1}. At large inter-nuclear distance, the cross term $\langle R|V_R|L\rangle$ decays very rapidly, approximately as $-E_{\text{Ha}} \exp(-R/a_0)\, a_0/R$. Thus for large R

$$\begin{aligned}\frac{\langle L|V_R|L\rangle + \langle R|V_R|L\rangle}{1 + \langle L|R\rangle} &= \langle L|V_R|L\rangle \frac{1 + \frac{\langle R|V_R|L\rangle}{\langle L|V_R|L\rangle}}{1 + \langle L|R\rangle} \\ &\simeq \langle L|V_R|L\rangle \left(1 + \frac{\langle R|V_R|L\rangle}{\langle L|V_R|L\rangle} - \langle L|R\rangle\right) \\ &\simeq \langle L|V_R|L\rangle \left(1 + \frac{\frac{-e^2}{R/2}\langle R|L\rangle + \cdots}{\frac{-e^2}{R}\langle L|L\rangle + \cdots} - \langle L|R\rangle\right) \\ &\simeq \langle L|V_R|L\rangle \left(1 + \langle L|R\rangle + \cdots\right), \end{aligned}$$

where the approximation $\langle R|V_R|L\rangle \simeq -\langle L|R\rangle e^2/(R/2)$ applies since the distribution $\psi_L(r)\psi_R(r)$ is very small everywhere, and peaks at the axial region between the two atoms, at the center of which $V_R \simeq -e^2/(R/2)$. As a consequence, for large R the attraction (last term of Eq. (3.13) for the symmetric state) *prevails over the repulsion* $V_{nn} = e^2/R$, thus generates *bonding*. The sign of the correction to the attraction is reversed for $|A\rangle$, where therefore the inter-nuclear repulsion prevails against attraction: the $|A\rangle$ state does not support bonding, hence its "antibonding" nature. This simple variational model predicts that, at large R, the adiabatic energy gain should be exponentially small in R/a_0, due to the decay of the 1s atomic wavefunctions.

All integrals in Eq. (3.12) can be evaluated for arbitrary R: Fig. 3.4 collects the outcome of this variational calculation. Note that:

- the adiabatic potential $V_{\text{ad}}^{(S)}(R)$ associated to the $|S\rangle$ state shows a minimum at a finite separation R_{M}; this potential well can bind the two protons together, and for this reason $|S\rangle$ is called a *bonding orbital*;
- the adiabatic potential $V_{\text{ad}}^{(A)}(R)$ is monotonically decreasing (repulsive), which suggests calling $|A\rangle$ an *antibonding orbital*;

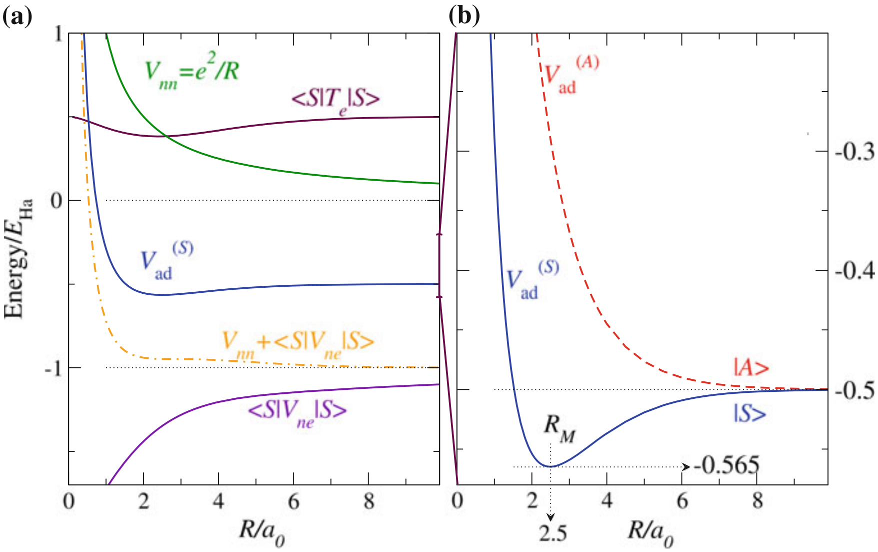


Fig. 3.4 **a** The total adiabatic potential of the $|S\rangle$ variational state of H_2^+ decomposed in its kinetic, potential and ion-ion contributions: $V_{ad}(R) = \langle S|T_e|S\rangle + \langle S|V_{ne}(r, R)|S\rangle + V_{nn}(R)$. **b** A blowup of the total adiabatic potential $V_{ad}(R)$ for the $|S\rangle$ (*solid*) and $|A\rangle$ (*dashed*) state. The minimum R_M of $V_{ad}(R)$ is indicated and corresponds to a molecular bond energy of 0.065 E_{Ha} or 1.76 eV

- the dot-dashed curve indicates that as R is reduced the lowering of the electronic potential energy $\langle S|V_{ne}(r, R)|S\rangle$ does not compensate the raise in inter-nuclear repulsion, thus our initial expectation that bonding is related to a gain in electrostatic potential energy is not confirmed by the present variational model;
- as the electron moves in the wider potential well created by the two protons rather than in the narrower well of an isolated proton, its kinetic energy decreases significantly, enough to dig a minimum in $V_{ad}^{(S)}(R)$.

The conclusion of our simple variational model for H_2^+ is that the physical origin of bonding is to be attributed to both kinetic and potential energy lowering of the electron "screening" the nuclei and spending a significant fraction of probability in the inter-nuclear region. This mechanism works optimally at $R \simeq 2.5\, a_0$, and generates a molecular bond energy of about 1.76 eV. For smaller R, the adiabatic potential shoots up due to the divergence in V_{nn} (now compensated poorly), and a new increase of the kinetic energy, as the molecular potential well shrinks to a more atomic-like shape.The proposed variational estimate provides basic qualitative trends, but is especially inaccurate at small R: a more quantitative treatment requires solving the Schrödinger equation in the double Coulomb well: this yields a bond energy of 2.79 eV at an optimal distance of $2.00\, a_0$, in good agreement with experiment.

The symmetry of the Schrödinger problem for one electron in the field generated by two nuclei is axial (cylindrical), rather than spherical as in atoms. Accordingly, the adiabatic electronic states are labeled by their angular-momentum *projection* m along the molecular axis $\hat{\mathbf{z}}$ through the nuclei, and not by their total angular

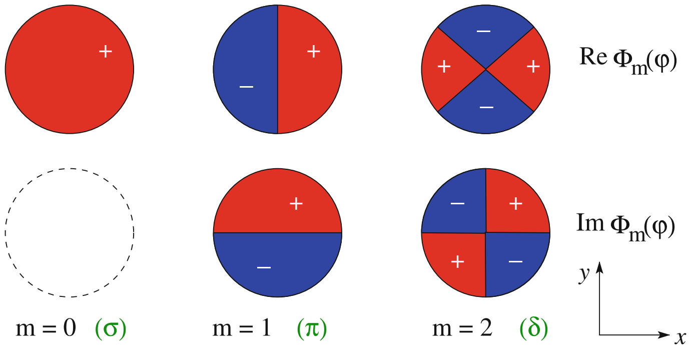


Fig. 3.5 The sign of the real and imaginary parts of the angular wavefunction of $m = 0, 1, 2$ electronic molecular states around the $\hat{\mathbf{z}}$ symmetry axis. Reverse-rotating $m = -1, -2$ states (not drawn) have $\mathrm{Im}\Phi_m(\varphi)$ of opposite sign

momentum l as in atoms. The electronic states of linear molecules, sketched in Fig. 3.5, are labeled σ, π, δ,..., to indicate the absolute value $|m| = 0, 1, 2, \ldots$ of their angular-momentum projection. Non-σ states are twofold degenerate orbitally. Both $|S\rangle$ and $|A\rangle$ of Eq. (3.11) are σ states, as being composed of 1s ($l = 0$, thus $m = 0$) atomic states. Standard chemical notation adds a star apex to antibonding states, thus $|A\rangle$ would in fact be labeled $1\sigma^*$. H_2^+ has further excited (bonding or antibonding) electronic states (derived from p, d,... atomic levels) which possess axial angular momentum $m \neq 0$, thus symmetry other than σ.

3.2.2 *Covalent and Ionic Bonding*

The model for the H_2^+ ion is instructive because in its simplicity it illustrates the physical origin of chemical bonding, sketches the basic concepts of bonding and antibonding orbitals, and provides a numerical estimate of typical bond energies and distances in molecules. Nonetheless, the electronic ground state of H_2^+ does not qualify as a proper chemical bond. For a full *covalent bond*, it would take *two electrons occupying* (clearly, with antiparallel spins) *the same bonding orbital* $|S\rangle$, as in neutral H_2. The precise one-electron molecular orbital $|S\rangle$ should be determined by some method (e.g. HF) which accounts—at least approximately—for electron-electron repulsion; eventually, $|S\rangle$ will turn out similar to our variational guess, Eq. (3.11). Accordingly, the ground wave function of a neutral H_2 molecule is described approximately by

$$\begin{aligned}\psi_e^{\mathrm{GS}}(r_1\sigma_1, r_2\sigma_2) &= \langle r_1\sigma_1, r_2\sigma_2|S\uparrow, S\downarrow\rangle^A \\ &= \langle r_1|S\rangle\langle r_2|S\rangle\, \frac{1}{\sqrt{2}} \begin{vmatrix} \chi_\uparrow(\sigma_1) & \chi_\uparrow(\sigma_2) \\ \chi_\downarrow(\sigma_1) & \chi_\downarrow(\sigma_2) \end{vmatrix}. \end{aligned} \tag{3.14}$$

Both electrons occupy the same *spatially symmetric* bonding orbital, forming a 2-electron state which is *antisymmetric* for exchange of electron 1 and 2 (through its spin-singlet part). The wavefunction (3.14) represents a typical *homonuclear covalent bond*. For H_2, the bond energy $V_{\mathrm{ad}}^{(\mathrm{GS})}(\infty) - V_{\mathrm{ad}}^{(\mathrm{GS})}(R_{\mathrm{M}}) \simeq 4.8\,\mathrm{eV}$ (at an equilibrium distance $R_{\mathrm{M}} \simeq 1.4\,a_0 = 74\,\mathrm{pm}$). This amount is less than twice the bond energy of H_2^+, due to the significant electron-electron repulsion payed by state (3.14).

The potential well generated by the two protons has room for several excited states, which are often single-electron excitations with one electron promoted to a higher bonding/antibonding state. Excited states with both electrons in bonding orbitals, also lead to a bonding $V_{\mathrm{ad}}(R_{\mathrm{M}})$, although usually characterized by lower binding energy and longer interatomic equilibrium distance than the electronic ground state [24]. Not all excited electronic states need to favor bonding: for example, the lowest triplet state of H_2, with electron one in $|S\rangle$ and electron two in $|A\rangle$ is not bound.

When pairs of larger-Z atoms interact to form dimers, to accommodate all electrons, the number of involved bonding and antibonding orbitals is necessarily larger. Figure 3.6 sketches a qualitative ground-state electronic structure of several homonuclear diatoms, leaving the core 1s electrons out. This figure prompts several important observations:

- With few exceptions, the overall ordering of the one-electron levels does not change much in passing from one molecule to another. In particular every antibonding level stands higher up in energy than the corresponding bonding level.

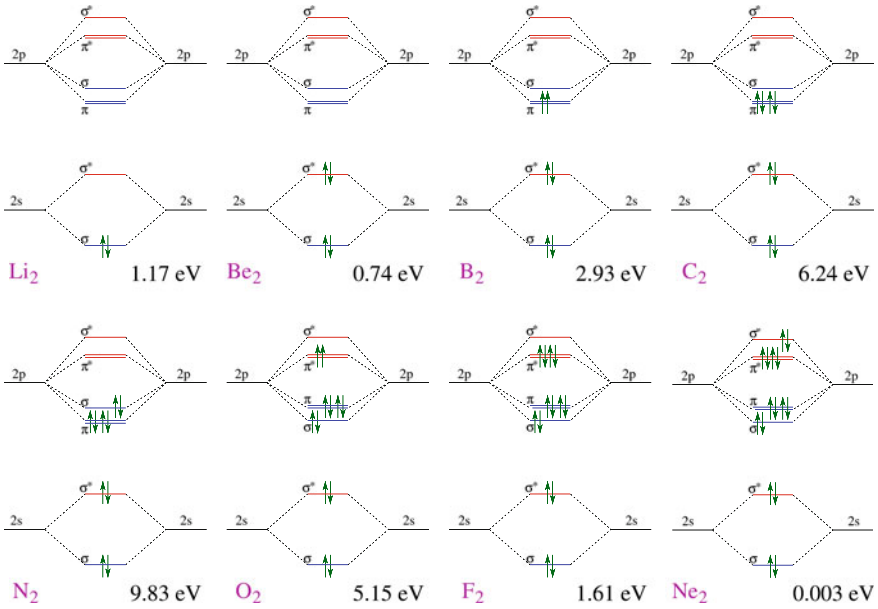


Fig. 3.6 A qualitative level scheme of the electronic structure of a few simple homonuclear diatomic molecules, with their observed bond energy. Bonding/antibonding single-electron orbitals are identified by their color. *Arrows* represent electrons occupying the orbitals

- Electrons fill the levels according to the same scheme as in atoms (low to higher energy, fulfilling Pauli's principle).
- The *bond order* is the number of electron *pairs* in a bonding level and missing in the corresponding antibonding level: e.g. O_2 is double-bonded.
- The strength of covalent bonds increases with their bond order: among the dimers of Fig. 3.6, the strongest bond is the triple bond of N_2, with a bond energy nearing 10 eV.
- When electrons start to fill 2p-derived molecular orbitals, the π states are lower than the σ combination. In a way, this is surprising, as the overlap of the $m = 0$ p orbitals (pointing along the molecular axis) is much larger than that of the $m = \pm 1$ orbitals (mostly located away from the axis): the $\sigma - \sigma^*$ splitting is generally larger than that of the corresponding $\pi - \pi^*$ orbitals. However, the Coulomb repulsion of the filled 1s-derived orbitals and the hybridization with the 2s-derived orbitals pushes the 2p-derived σ orbitals up by a substantial amount. The "natural" ordering is restored in O_2.
- In B_2 and O_2, two electrons sit in a degenerate π or π^* molecular orbital, which has room for 4 electrons: to minimize the residual Coulomb repulsion, the ground state is a spin-triplet Hund-rules state $\begin{array}{|c|c|}\hline \uparrow & \uparrow \\ \hline & \\ \hline \end{array}$ (Sect. 2.2.8.3). These dimers are magnetic, and their gases are therefore paramagnetic (see Sect. 4.3.1.3).
- Even dimers such as Be_2 and Ne_2, with null bond order (equally populated bonding and antibonding orbitals), exhibit a weak bonding, of the type described in Sect. 3.2.3.

Homonuclear diatomic molecules are very special diatoms, due to their peculiar $L \leftrightarrow R$ symmetry. Most pairs of atoms bind together, and many form covalent bonds not so unlike those illustrated for the homonuclear molecules. For example, the CO molecule has the same number of electrons as the molecule N_2. The main novelty is the lack of $L \leftrightarrow R$ symmetry: the attraction of the O nucleus ($Z = 8$) is stronger than that of C ($Z = 6$). As a consequence, as sketched in Fig. 2.21, the 2s and 2p shells of O sit deeper than the 2s and 2p shells of C: the bonding and antibonding orbitals deviate from symmetric/antisymmetric combinations of the type (3.11). Extending the variational treatment of H_2^+ with the methods of Appendix B.5, with the simplification of neglecting overlaps $\langle L|R\rangle$, the approximate molecular orbitals are the eigenstates of the 2×2 matrix

$$\begin{pmatrix} E_L & -\Delta \\ -\Delta & E_R \end{pmatrix} \equiv \begin{pmatrix} \langle L|H_1|L\rangle & \langle L|H_1|R\rangle \\ \langle R|H_1|L\rangle & \langle R|H_1|R\rangle \end{pmatrix}, \tag{3.15}$$

where $H_1 = T_e + V^{\text{eff}}$ is the effective one-electron Hamiltonian fixed by the self-consistent potential acting on each electron, and $|L\rangle$, $|R\rangle$ are the considered atomic states (e.g. the 2s or 2p) of the left and right atom (here C and O). The diagonal elements are mostly dictated by the energy positions of atomic shells. The off-diagonal overlap energy $\Delta > 0$ is very small at large separation, and grows larger and larger as the atoms move closer.

The eigenvalues $\mathscr{E}_b = \bar{E}_1$ and $\mathscr{E}_a = \bar{E}_2$ of the 2×2 matrix (3.15) are discussed in Appendix B.5.2, and shown in Fig. B.2. The eigenenergies in Eq. (B.39) remain centered around $(E_L + E_R)/2$, but their splitting $\mathscr{E}_a - \mathscr{E}_b = \left[(E_L - E_R)^2 + (2\Delta)^2\right]^{1/2}$ exceeds both the diagonal splitting $E_L - E_R$ and the minimum splitting $|2\Delta|$, which is recovered in the symmetric homonuclear limit $E_L = E_R$ (i.e. $u = 0$). In this symmetric limit the bonding state $|b\rangle = |\bar{E}_1\rangle$ and antibonding state $|a\rangle = |\bar{E}_2\rangle$ coincide with the equally-weighted $|S\rangle$ and $|A\rangle$ combinations of Eq. (3.11). For increasing u, the eigenkets given in Eq. (B.40) resemble less and less the simple symmetric and antisymmetric combinations (3.11): $|b\rangle$ acquires a prevalent $|R\rangle$ character, while $|a\rangle$ acquires a mainly $|L\rangle$ character. Electrons in the $|b\rangle$ state move partly to the R side, thus providing a *polar character* to the dimer (unless an equal number of electrons occupies the $|a\rangle$ state).

As sketched in Fig. 3.7a, an intermediate value of $u \simeq 1$ applies for the 2s and 2p orbitals of CO near its equilibrium separation: bonding molecular orbitals lie prevalently at the O side, antibonding ones at the C side. The bond of CO and of many similar heteronuclear diatoms is classified as *polar covalent*, since it is associated to a nonzero average electric dipole due to the charge transfer produced by the asymmetric charge distribution of the electrons in the bonding state $|b\rangle$.

In the limit of very large polar character ($u \gg 1$), the eigenenergies (B.39) $\mathscr{E}_{\substack{b\\a}} \simeq E_{\substack{R\\L}}$, and the eigenkets $|b\rangle \simeq |R\rangle$, $|a\rangle \simeq |L\rangle$. At typical interatomic separation, the bond of HF (here hydrogen fluoride, not Hartree-Fock!) is characterized by large u, see Fig. 3.7b. Recall—Fig. 2.21—that $\mathscr{E}_{1s}(\mathrm{H}) \ll \mathscr{E}_{2p}(\mathrm{F})$. As a result, the relevant bonding orbital is not much different from the 2p of an isolated F atom. Thus an approximate description of the bond of HF invokes a complete charge transfer from H to F: the latter therefore completes its open shell to $2p^6$. As long as the H^+ and F^- ions are separated widely, they would then attract each other like point charges, with a $-e^2/R$ attraction energy. As the proton moves inside the outer shell of F^-, electrons screen the positive charge of the F nucleus less and less, and this attraction

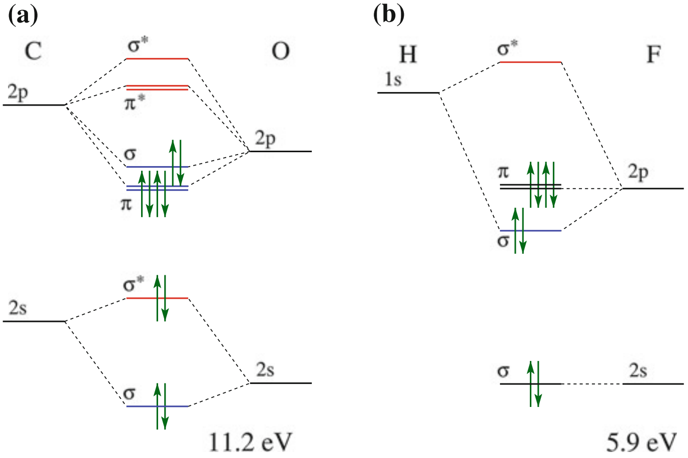


Fig. 3.7 A schematic electronic structure of two heteronuclear diatomic molecules: **a** CO and **b** HF. *Blue* indicates bonding, *red* indicates antibonding single-electron orbitals. The observed bond energy is indicated. Single-electron energies and splittings are purely qualitative

turns gradually into a repulsion. The origin of bonding in this simple picture is the energy gained in moving the ions from infinity to the equilibrium distance, to which the energy paid to form the ions from the neutral atoms (the ionization potential of the atom which turns into a cation—here H—minus the electron affinity of the atom acquiring the electron—here F) is to be subtracted. Indeed, $[e^2/(92\,\text{pm}) = 15.7\,\text{eV}]$ minus the ionization potential of H (13.6 eV) plus the electron affinity of F (3.4 eV) yields 5.5 eV, in fair agreement with the observed bond energy $= 5.9\,\text{eV}$. A molecular bond such as that of HF, where u is so large that almost complete charge transfer occurs is named an *ionic bond*.

The picture of molecular orbitals constructed by specific atomic orbitals, as in Figs. 3.6 and 3.7 and relative discussion, is oversimplified. In fact, all orbitals of the same symmetry are mixed, although with one prominent atomic component in the actual molecular orbital. For example, each one of the orbitals labeled 2σ $2\sigma^*$ 3σ and $4\sigma^*$ in Fig. 3.6 is a linear combination involving a little of all $m = 0$ states of both atoms, mostly 2s and $2p_z$, but also 1s, 3s, $3p_z$, 4s… This *hybrid* character of molecular orbitals, a side detail for diatomic molecules, is a crucial ingredient in determining the 3D structure of *polyatomic molecules* and covalent solids.

Hybrid orbitals generate adiabatic potentials which depend strongly not only on interatomic distances, but also on angles. An especially remarkable example is the mix of 2s and 2p, at the origin of the shape of organic molecules, Fig. 3.8. As specific and important examples, sp^3 combinations intermixing 2s $2p_x$ $2p_y$ and $2p_z$ orbitals determine the ideally 109° angle between the bonds of tetrahedrally coordinated carbon and silicon; sp^2 combinations of 2s $2p_x$ $2p_y$ tend to bind atoms such as carbon to three other atoms in the same plane, forming ideally 120° angles. The endless combinations of the molecular orbitals of polyatomic 3D molecular structures, plus the multitude of weaker interactions among individual molecules, originate the infinite variety of organic compounds, which are the microscopic building blocks and byproducts of living matter. Analogous mixtures of 3s 3p, or 4s 4p, provide Si, P, S, Ge, As, Se with a marked tendency to form directional bonds, producing the extended (rather than molecular as for N_2) covalently bound extended structures characterizing the solid state of these elements (see Figs. 5.9, 5.25, and 5.62). Each equilibrium molecular structure identifies a minimum of the adiabatic potential V_{ad}. Polyatomic molecules are often characterized by multiple local minima of V_{ad}, representing individual *isomers*, e.g. those of Fig. 3.8i–k.

3.2.3 Weak Non-chemical Bonds

Our description of covalent bonds in terms of linear combinations of single "rigid" atomic orbitals, is doomed to obtain an exponential dependence of $V_{\text{ad}}(R)$ at large distance R, due to the exponential decay of the orbitals themselves. Experimentally however, at large distance, the interaction energy between any two atoms is attractive and decays following a universal power law: $\sim R^{-6}$. This is due to a

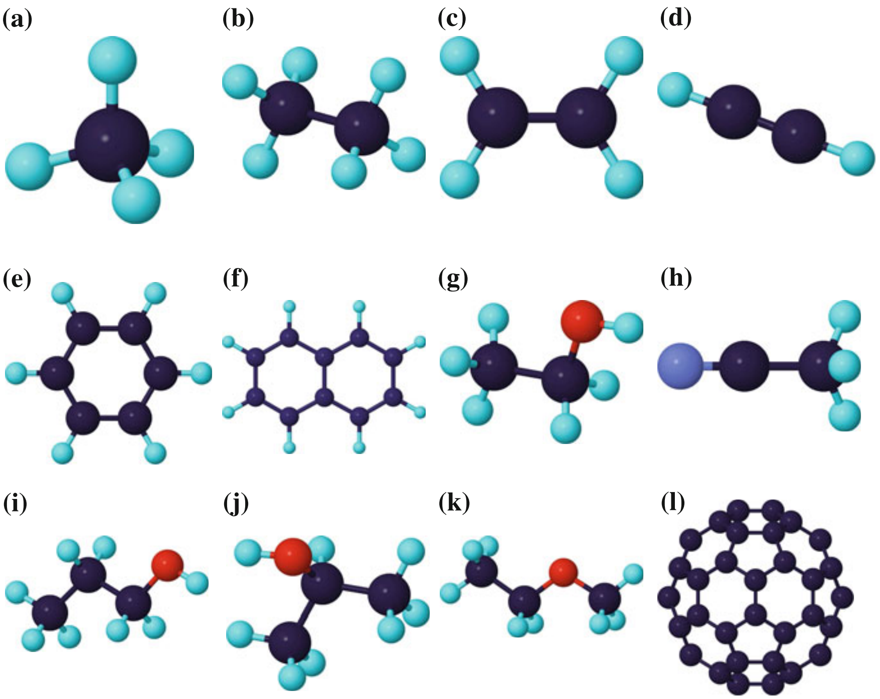


Fig. 3.8 The 3D equilibrium structures [25] of organic molecules is dictated by the hybrid orbitals of carbon. sp^3 hybridization, with C atoms at the center of a tetrahedron of 4 bonds, is relevant e.g. for **a** methane CH_4, **b** ethane C_2H_6, **g** ethanol CH_3CH_2OH, the carbon in the CH_3 group of **h** acetonitrile, and **i**–**k** three isomers of C_3H_8O. sp^2 hybridization, with C atoms at the center of a triangle of 3 bonds, is relevant e.g. for **c** ethylene C_2H_4, **e** benzene C_6H_6, and **f** naphthalene $C_{10}H_8$. sp hybridization with C atoms in a linear 2-bonds configurations is relevant for **d** acetylene C_2H_2, and the carbon in the CN group of **h** acetonitrile. The hybridization of **l** fullerene C_{60} is intermediate between sp^2 and sp^3

classical-electromagnetism concept which the previous analysis completely overlooked: atomic polarizability.

On average, the electric dipole moment of any atom vanishes. As long as spherical symmetry is unbroken, the single-electron atomic orbitals are also eigenstates of the orbital angular momentum $|\mathbf{l}|^2$. The parity of the eigenfunctions $R(r) \times Y_{l\,m_l}(\hat{\mathbf{r}})$ is the same as that of its spherical harmonic: $(-1)^l$; thus the angular probability distribution $|Y_{l\,m_l}(\hat{\mathbf{r}})|^2$ of each electron is parity-even (see also Fig. 2.5 and Sect. 2.1.9). As a result, the average electric dipole $\langle l\,m_l|\mathbf{d}|l\,m_l\rangle$ vanishes. However, electrons move around the nucleus and occupy instantaneously specific positions, associated to a fluctuating nonzero dipole moment. According to basic electromagnetism, this dipole moment generates an instantaneous electric field $\mathbf{E}$, whose intensity decays away from the atom as R^{-3}.

An electric field acting on a second remote atom "polarizes" it. The electrons of atom b react to minimize the total energy, now including an additional term $-\mathbf{d} \cdot \mathbf{E}$

describing the coupling to the external field, where $\mathbf{d} = -q_e \sum \mathbf{r}_i$. Atom b responds to the applied external electric field by building up an *induced dipole* in the same direction as the field: $\langle \mathbf{d} \rangle = \alpha_b \mathbf{E}$. As the field is weak, the *atomic polarizability* α_b is independent of the field (*linear response*). Due to this induced polarization, the total energy of the two atoms at separation R decreases by an amount $-1/2 \langle \mathbf{d} \rangle \cdot \mathbf{E} = -1/2 \alpha_b |\mathbf{E}|^2 \propto R^{-6}$ compared to atoms a and b placed at infinite separation. Dimension-wise, the coefficient multiplying R^{-6} is a [distance]^6 × [energy]. As only atomic physics is involved, the order of magnitude of this coefficient is $E_{\mathrm{Ha}}\, a_0^6 \simeq 10^{-79}\,\mathrm{J\,m^6} \simeq 0.6\,\mathrm{eV\,Å^6}$. This *van der Waals attraction* is then approximately

$$V_{\mathrm{ad}}^{\mathrm{vdW}}(R) \approx -E_{\mathrm{Ha}} \left(\frac{a_0}{R}\right)^6 . \tag{3.16}$$

More quantitatively, the coefficient of R^{-6} is proportional to the product of the two atomic electrical polarizabilities $\alpha_a \alpha_b$ (because the dipole fluctuation in atom a is also proportional to its polarizability α_a), or, more precisely, to the energy integral of their energy-dependent version $\alpha_{a/b}(\hbar\omega)$ [26]. The physical dimension of polarizability is $[\mathrm{Charge}]^2[\mathrm{Length}]^2[\mathrm{Energy}]^{-1} = [\mathrm{Charge}]^2[\mathrm{Time}]^2[\mathrm{Mass}]^{-1}$, the same as $4\pi\varepsilon_0 \times [\mathrm{Length}]^3$. It is therefore natural to express atomic polarizabilities in units of $\alpha_{\mathrm{A.U.}} = 4\pi\varepsilon_0\, a_0^3 = 1.64878 \times 10^{-41}\,\mathrm{C^2 s^2/kg}$.

The atomic polarizability can be either measured (for example, by measuring the dielectric properties of an atomic gas inserted in a charged parallel-plate capacitor) or computed, by evaluating the distortion of the atomic state under the action of an applied electric field. By standard perturbation theory, Appendix B.9, an external field distorts the ground state so that (i) the electronic wavefunction acquires components (proportional to $|\mathbf{E}|$) of excited states with l changed by ± 1, and (ii) the total energy lowers quadratically with the field. For example, in a weak external electric field $\mathbf{E} = E_z \hat{\mathbf{z}}$, the ground state of an H atom takes the form of a linear combination

$$|1, 0, 0; E_z\rangle = b_1 |1, 0, 0\rangle + \sum_{n>1} b_n |n, 1, 0\rangle . \tag{3.17}$$

Here the hydrogen-atom eigenkets $|n, l, m\rangle$ are represented by the eigenfunctions detailed in Eq. (2.1.6). The acquired electric dipole moment $\langle 1, 0, 0; E_z | d_z | 1, 0, 0; E_z \rangle = 2 b_1 \sum_n b_n \langle n, 1, 0 | d_z | 1, 0, 0 \rangle$ is proportional to E_z, because the $b_{n>1}$ coefficients are. To evaluate the polarizability α, estimate the coefficients $b_{n>1}$ by first-order perturbation theory, Eq. (B.87), $b_n = -E_z \langle n10 | d_z | 100 \rangle / (\mathscr{E}_1 - \mathscr{E}_n)$, where $\mathscr{E}_n$ are the energy levels of H, Eq. (2.10). The atomic polarizability α is then given by

$$\alpha = \frac{\langle 1, 0, 0; E_z | d_z | 1, 0, 0; E_z \rangle}{E_z} \simeq -2 \sum_{n>1} \frac{|\langle n, 1, 0 | d_z | 1, 0, 0 \rangle|^2}{\mathscr{E}_1 - \mathscr{E}_n} . \tag{3.18}$$

The static atomic polarizability of H amounts to $\alpha_{\mathrm{H}} = 4.50\, \alpha_{\mathrm{A.U.}}$. Taking all electrons into account, one can evaluate the same quantity for any atom. For example, He has a record small $\alpha_{\mathrm{He}} = 1.383\, \alpha_{\mathrm{A.U.}}$ due to its huge 1s-2p gap; Li has a much larger

$\alpha_{\rm Li} = 164.0\,\alpha_{\rm A.U.}$ due to the comparably small 2s-2p gap; and Cs holds the record $\alpha_{\rm Cs} = 401.0\,\alpha_{\rm A.U.}$ mostly due to the very small 6s-6p gap. Indeed, Eq. (3.18) suggests that α depends inversely on the energy separation between the highest filled state and the lowest empty state with l differing by ± 1.

The dipole–induced-dipole mechanism is perfectly general, and it accounts for the leading long-range attraction of all dimers of neutral atoms. As distance R decreases, the dimer electrons can evolve along two distinct paths: open-shell atoms, e.g. H and those forming the diatoms of Figs. 3.6 and 3.7, modify gradually their orbitals to form robust covalent bonds; in contrast, for pairs of close-shells atoms (Be_2, Ne_2) the one mechanism producing a (comparably weak) attraction is van der Waals, until short-distance repulsion turns in.

As R is further reduced, the short-distance repulsive Coulombic divergence of $V_{\rm ad} \propto R^{-1}$ is peculiar to H_2^+ and few other dimers involving H. For general many-electron atoms, long before the nucleus-nucleus repulsion $V_{nn}(R)$ becomes relevant, the electronic energy $E_e^{(a)}(R)$ blows up because of Pauli's principle: as the core electrons are brought together in the same region of space, their wavefunctions become less and less orthogonal. Part of these electrons are then pushed up into some orthogonal empty valence level, which makes tiny reductions of R cost tens or even hundreds $E_{\rm Ha}$. This rapid "hard-core" energy increase as two atoms collide is responsible for the "impenetrability" of matter, i.e. the sharp increase of pressure of a sample whose volume is reduced so much that each atom is squeezed into less than its characteristic volume ($\sim$ a few Å^3). We see that a combined effect of electrons indistinguishableness and quantum kinetic energy sustains matter against collapse due to electromagnetic attraction. In a diatomic context, the $V_{\rm ad}(R)$ blowup at short distance is often parameterized phenomenologically with a R^{-12} power law.

The simplest potential capturing the long-distance van der Waals attraction and the rapid short-range repulsion is the popular Lennard-Jones potential

$$V_{\rm LJ}(R) = 4\varepsilon \left[\left(\frac{\sigma}{R}\right)^{12} - \left(\frac{\sigma}{R}\right)^{6} \right]. \tag{3.19}$$

This expression is nothing but a phenomenological model for the actual $V_{\rm ad}(R)$ of a dimer. Its two parameters ε (the depth of the potential well at $R_{\rm M}$) and σ (the radius where $V_{\rm LJ}(R)$ changes sign) are listed in Table 3.1 for the noble-gas

Table 3.1 Parameters (obtained with a fit to atom-atom scattering data [2]) for the Lennard-Jones pair potentials, Eq. (3.19), of the dimers of the noble-gas elements. Note the increase of atomic size (i.e. σ) with Z and the rapid increase in atomic polarizability in going from He to Ne [reflected by the coefficient $4\varepsilon\sigma^6 \propto \alpha^2$ of $-R^{-6}$ in Eq. (3.16)]

Element	σ (pm)	ε (meV)	$4\varepsilon\sigma^6$ ($E_{\rm Ha}\,a_0^6$)
He	256	0.879	1.7
Ne	275	3.08	8.9
Ar	340	10.5	109
Kr	368	14.4	239
Xe	407	19.4	590

dimers, whose actual $V_{\text{ad}}(R)$ the Lennard-Jones potential is a fair approximation for. In this context, $V_{\text{LJ}}(R)$ is also used for describing the dynamics of a collection of more than two noble-gas atoms as a *sum of pair potentials* (*two-body* forces). This extrapolation is a fair approximation of the actual adiabatic potential of simple close-shell systems at low density: the phase diagram and correlation properties of the Lennard-Jones solid/fluid is in close qualitative and semi-quantitative agreement to observed data for noble-gas systems. The two-body Lennard-Jones model instead is highly inappropriate for atoms forming strongly directional covalent bonds.

3.2.4 Classification of Bonding

We have clarified the mechanism for the general tendency, pictured in Fig. 3.1, of atoms to attract at large distance and repel when coming in contact. Contrasting this qualitative likeness, significant differences in the equilibrium distances and huge differences in well depths make different diatoms, bonded by different mechanisms, very unlike.

When at least one of the atoms is a noble gas, the dipole–induced-dipole mechanism is the only mechanism creating attraction, little or no covalency occurs, and consequently the equilibrium atom-atom distance R_{M} is rather large ($\sim$250–400 pm) and the bond energy $[V_{\text{ad}}(+\infty) - V_{\text{ad}}(R_{\text{M}})]$ is small (few meV, see Table 3.1). Weakly bonded van der Waals systems retain ordinarily the monoatomic gas phase down to relatively low temperature, and show a scarcer tendency to form dimers than to condense to extended liquid and solid phases.

A few atoms with open shells (O, N, F) display a prominent tendency to form diatomic molecules, with short strong covalent bonds, characterized by typical several-eV bond energies. These molecular units are retained in the low-temperature liquid and solid phases. For most other elements (e.g. Li, Be, B, C,...), the extra energy gain in forming multiple chemical bonds per atom makes it energetically convenient to form extended metallic or covalent solids, rather than diatoms.

When different atoms are bound together covalently, a fraction of their valence charge moves closer to certain ions than to others, because the energy of the involved atomic shells is different, as illustrated in the examples of Fig. 3.7. This electron displacement provides a nonzero average electric dipole to *polar* heteroatomic bonds. Even in the extreme limit of complete or almost complete charge transfer (e.g. HF, LiF), energies and lengths involved in *ionic bonds* are in the same range as for covalent bonds.

3.3 Intramolecular Dynamics and Spectra

In the previous section, we have discussed general properties of the solutions of the electronic equation (3.2) for a diatom, thus acquiring information on the typical shape of the adiabatic potential V_{ad}. We consider now the motion of the two nuclei in the adiabatic force field described by V_{ad}, with its spectroscopic implications. This motion can be described in terms of solutions of Eq. (3.9).

As remarked in Sect. 3.1, the adiabatic potential is independent of the center-mass position of the molecule (translational invariance) and of the orientation in space of the straight line through the two nuclei (rotational invariance). Precisely the same transformation (2.1) applied to the two-body problem of the one-electron atom separates the center-mass motion of the two-body problem of the diatomic molecule. Like for atoms, due to translational symmetry, the molecular center of mass translates freely: the random thermal translational motion in a gas-phase sample originates Doppler and collisional broadening of the spectra.

By using polar coordinates for the vector joining the two nuclei, we separate the Schrödinger equation for the internal motion into angular (2.5), (2.6) and radial (2.7) equations, exactly like for the one-electron atom. Of course, in the context of diatoms, the nucleus-nucleus separation R replaces the electron-nucleus distance r of the one-electron atom and the distance-dependent adiabatic potential $V_{\mathrm{ad}}(R)$ replaces $U(r)$ in Eq. (2.7). The angular equations are universal, so that the angular eigenfunctions, describing the orientation in space of the molecule (thus molecular rotations) are standard spherical harmonics $Y_{l\,m_l}$. Rotational states $|l, m_l\rangle$ are labeled by the molecular angular momentum l,[1] plus its z component m_l.

The formal structure of the radial equation is the same as for the one-electron atom. The substantial physical difference stands in the equilibrium distance (the separation where the potential is the most attractive) which for the diatom is finite $R = R_{\mathrm{M}} > 0$, rather than $R = 0$ as for the one-electron atom. The main consequence is that the radial motion is mostly localized near R_{M}, in a region where the centrifugal term $\hbar^2 l(l+1)/(2\mu R^2)$ in Eq. (2.7) is often fairly small. If we neglect the variations of R^{-2} along a region around R_{M} where the radial wavefunctions differ significantly from zero, then the radial motion is approximately independent of the rotation, i.e., the radial solutions are independent of l. The radial quantum number $v = 0,\ 1,\ 2, \ldots$ for the diatomic molecule counts the radial nodes, matching $n-l-1$ in the one-electron atom. If the adiabatic potential is Taylor-expanded around its minimum

$$V_{\mathrm{ad}}(R) = V_{\mathrm{ad}}(R_{\mathrm{M}}) + \frac{1}{2}\left.\frac{d^2 V_{\mathrm{ad}}(R')}{dR'^2}\right|_{R'=R_{\mathrm{M}}} (R - R_{\mathrm{M}})^2 + \cdots\,, \tag{3.20}$$

and truncated at second order, then the radial motion is approximately a *harmonic* motion. In this scheme, the total energy consists of the sum of three contributions, of decreasing magnitude:

- A large *electronic* term $V_{\mathrm{ad}}(R_{\mathrm{M}})$, whose lowering at the equilibrium configuration $V_{\mathrm{ad}}(\infty) - V_{\mathrm{ad}}(R_{\mathrm{M}})$ measures the well depth responsible for the chemical bond.

---

[1] In the literature, the angular-momentum quantum number describing the overall molecular rotation is occasionally labeled r, or j, rather than l.

- A *vibrational* term, which in the harmonic approximation, Eq. (B.65), amounts to

$$\mathscr{E}_{\mathrm{vib}}(v) = \hbar\omega\left(v + \frac{1}{2}\right), \tag{3.21}$$

measuring the energy of radial vibration around the equilibrium position R_{M}. Here

$$\omega = \sqrt{\frac{K}{\mu}}, \quad \text{with } K = \left.\frac{d^2 V_{\mathrm{ad}}(R')}{dR'^2}\right|_{R'=R_{\mathrm{M}}} \quad \text{and } \mu = \frac{M_1\, M_2}{M_1 + M_2}, \tag{3.22}$$

where μ indicates the reduced mass of the two-body oscillator composed of the two nuclei of mass M_1 and M_2 [see Eq. (2.4)]. Typical vibrational energies $\hbar\omega$ of few hundred meV or less are observed.

- A rotational contribution from the mean rotational term $\hbar^2 l(l+1)/(2\mu R^2) \simeq \hbar^2 l(l+1)/(2\mu R_{\mathrm{M}}^2)$ in the radial equation (2.7) yields:

$$\mathscr{E}_{\mathrm{rot}}(l) = \frac{\hbar^2 l(l+1)}{2\mu R_{\mathrm{M}}^2} = \frac{|\mathbf{L}|^2}{2\,I}, \tag{3.23}$$

where I stands for the classical momentum of inertia μR_{M}^2 of the diatom relative to its center of mass, assuming the interatomic distance R is frozen at the equilibrium separation R_{M}. The typical order of magnitude of the rotational energies $\hbar^2/(2I)$ in molecular spectra is few meV or less, H_2 exhibiting the largest one: 7 meV.

3.3.1 Rotational and Rovibrational Spectra

In an "adiabatic" transition with the electrons remaining in the electronic ground state, observed spectra fulfill the standard dipole selection rules: $\Delta l = \pm 1$. The involved electric dipole operator is the product of the inter-nuclear separation times the charge difference permanently attached to the two atoms. This charge difference vanishes for equal nuclei, thus no dipole transition occurs for homonuclear molecules.[2] In contrast, the large permanent dipole moments of strongly polar molecules, such as HF and HCl, support intense dipole transitions. The dipole moment of CO, a weakly polar molecule, is only about 10 % that of HCl, thus producing much weaker IR absorption.

---

[2] This is the reason why clean air (composed mostly of gas-phase N_2 and O_2) is highly transparent to IR light. Transparency in the visible and near-UV range is associated to the wide gaps (several eV) from the electronic ground state to the first allowed electronic excitation.

Rotational and vibrational molecular spectra are mostly observed in *absorption* rather than emission. The reason is the small spontaneous emission rate of such low-energy transitions, due to the $\mathscr{E}^3$ dependence of the electric-dipole decay rate, Eq. (2.45). Other, radiationless, decay phenomena (typically related to molecular collisions) occur over far shorted times scales, thus making it unfeasible to observe emission spectra in this spectral region.

Purely rotational spectra, usually in the far IR region, are associated to $\Delta v = 0$, $\Delta l = 1$ transitions. The energy difference between the $|l_i\rangle$ and $|l_f\rangle = |l_i + 1\rangle$ states is

$$\Delta\mathscr{E}_{\text{rot}}(l_i) = \frac{\hbar^2}{2I}\left[(l_i+1)(l_i+1+1) - l_i(l_i+1)\right] = \frac{\hbar^2}{I}(l_i+1)\,. \tag{3.24}$$

Accordingly, the rotational spectrum of a sample composed of molecules in several initial rotational states is constituted by an array of equally spaced lines. The energy spacing is twice as large as the typical rotational energy quantum $\hbar^2/(2I)$. Figure 3.9 reports a characteristic purely rotational absorption spectrum. The measured separation of the lines permits us to estimate the interatomic equilibrium separation R_{M} through Eq. (3.24).

Rovibrational spectra are observed typically in near-IR absorption. Most of the dipole intensity concentrates in $\Delta v = 1$ transitions, although weaker *overtone* transitions with $\Delta v > 1$ are observed routinely. Figure 3.10 reports a characteristic rovibrational spectrum. Again, transitions occur starting from several initial rotational states: as a consequence, the purely vibrational peak is "decorated" by rotational transitions, which on the low-energy side imply $\Delta l = -1$ and are called P branch, and on the high-energy side imply $\Delta l = +1$ and are called R branch, as illustrated in Fig. 3.11. The rotational structures are equally spaced according to Eq. (3.24) (for the P branch a similar result holds). The rovibrational spectra are characteristic for the absence of a purely vibrational peak at energy $\hbar\omega$, with could only occur if $\Delta l = 0$ transitions were dipole allowed (which are not).

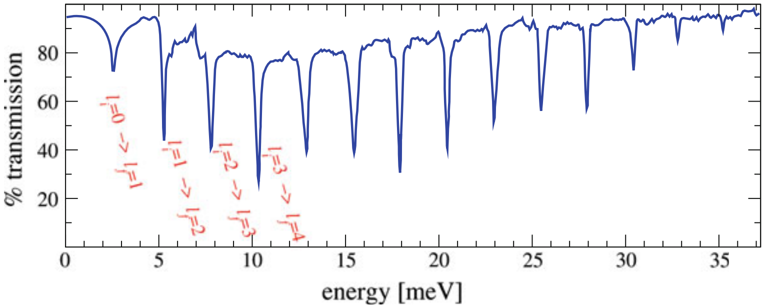


Fig. 3.9 Observed purely rotational spectrum of HCl gas, with a few $|l_i\rangle \to |l_f\rangle$ assignments. In this spectrum, the $|l_i = 0\rangle \to |l_f = 1\rangle$ line at ~2.5 meV is just sketched rather than observed, due to range limitations of the spectrometer (Data from Ref. [3])

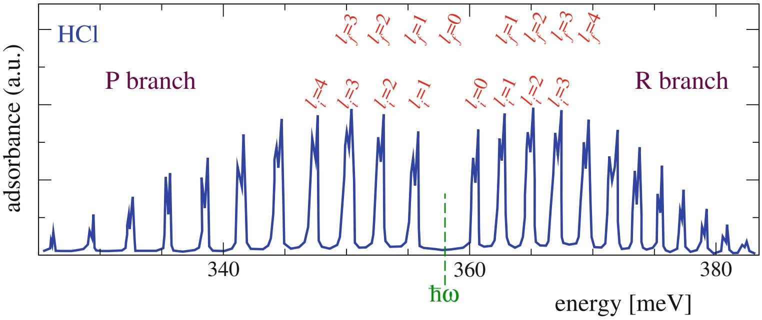
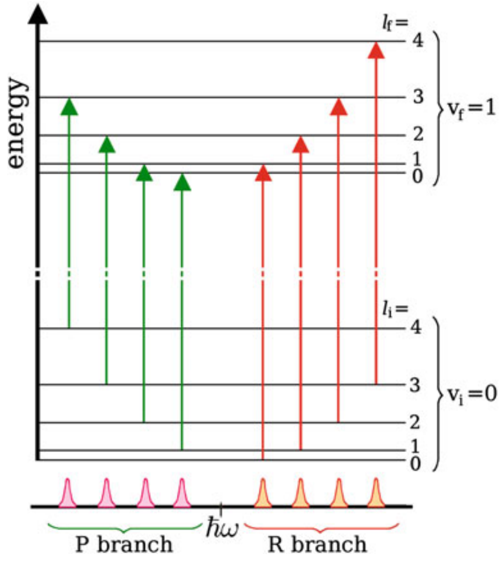


Fig. 3.10 Observed rovibrational spectrum of HCl gas: absorption in this region is associated to the "fundamental" transition $|v_i = 0\rangle \rightarrow |v_f = 1\rangle$, decorated by the rotational P branch transitions $|l_i\rangle \rightarrow |l_i - 1\rangle$, and R branch transitions $|l_i\rangle \rightarrow |l_i + 1\rangle$. The isotopic duplication of the lines is visible: $H^{35}Cl$ is responsible for the stronger peaks and less abundant $H^{37}Cl$ for the weaker ones

Fig. 3.11 A scheme of the rotational levels and the relative electric-dipole-allowed absorption transitions around the $v = 0 \rightarrow 1$ vibrational excitation (at energy $\hbar\omega$) of a diatomic molecule



Rovibrational spectra are often investigated through Raman spectroscopy, which is not based on dipole transitions as IR absorption. Raman experiments are based on optical (electronic) non-resonant excitations of the molecule, which rapidly decay back to the electronic ground state, possibly leaving a vibrational and/or rotational excitation. As the experiment involves two photons, the selection rules allow $\Delta l = 0$, ± 2 transitions.

The rovibrational dynamics of polyatomic molecules involves further intricacies. Once translations and rotations have been accounted for, $3N_n - 6$ (or $3N_n - 5$ for linear molecules) internal degrees of freedom account for vibrations, approximately

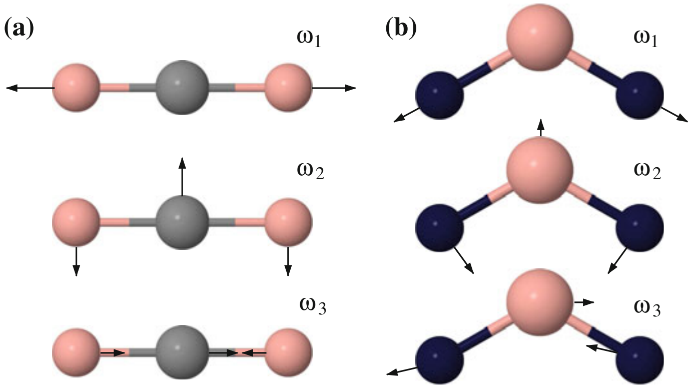


Fig. 3.12 The vibrational normal-mode displacements of two triatomic molecules. **a** CO_2, a molecule whose equilibrium geometry is linear. Its normal-mode energies are $\hbar\omega_1 = 166\,\text{meV}$, $\hbar\omega_2 = 83\,\text{meV}$, $\hbar\omega_3 = 291\,\text{meV}$. **b** H_2O, a bent molecule at equilibrium. Its normal-mode energies are $\hbar\omega_1 = 453\,\text{meV}$, $\hbar\omega_2 = 198\,\text{meV}$, $\hbar\omega_3 = 466\,\text{meV}$

described in terms of small-amplitude harmonic oscillations (the *normal modes* of classical mechanics) around the multi-dimensional minimum of V_{ad}. Figure 3.12 sketches such vibrational modes for CO_2 and H_2O. Each normal mode i behaves as an independent quantum harmonic oscillator, with its own vibrational frequency (and thus ladder step $\hbar\omega_i$).

3.3.2 Electronic Excitations

Visible and UV photons can excite electronic states. These transitions are similar to atomic excitations, and can roughly be understood as promotions of an electron from a filled molecular orbital to an empty one, e.g. from a bonding to an antibonding orbital. An electronic transition $\psi_e^{(a)} \to \psi_e^{(b)}$ in a molecule leads from one adiabatic potential surface to another, as illustrated in Fig. 3.13.

During the extremely short timescale of the electronic transition, the ions have no time to move, thus the transition is "vertical" in R. In general, the shapes of different adiabatic potentials $V_{\text{ad}}^{(a)}(R)$ and $V_{\text{ad}}^{(b)}(R)$ associated to different electronic states are often well distinct; in particular $R_{\text{M}}^{(b)} \neq R_{\text{M}}^{(a)}$. This means that electronic transitions are usually accompanied by vibrational "shakeup" transitions, excited by the displacement of the equilibrium geometry. For this reason, an electronic transition usually gets decorated by approximately equally-spaced vibrational satellites, separated by the harmonic frequency $\hbar\omega^{(b)}$ of the final adiabatic potential, as sketched in Fig. 3.14.

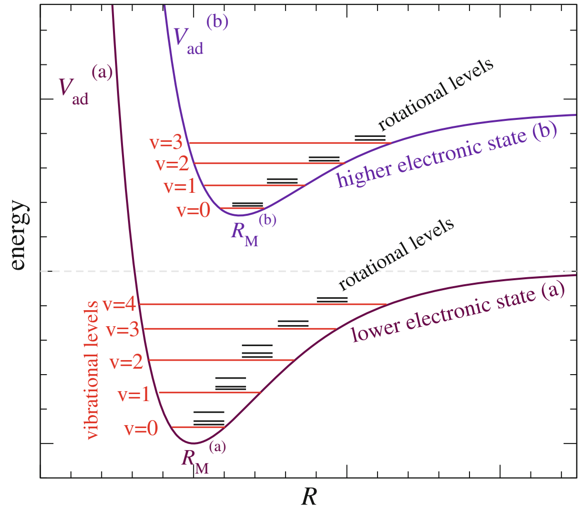


Fig. 3.13 Rotational states build up tiny nonuniformly-spaced ladders on top of each of the pure ($l = 0$) vibrational states of two adiabatic potentials $V_{\text{ad}}^{(a)}(R)$ and $V_{\text{ad}}^{(b)}(R)$ associated to the electronic eigenfunctions $\psi_e^{(a)}$ and $\psi_e^{(b)}$

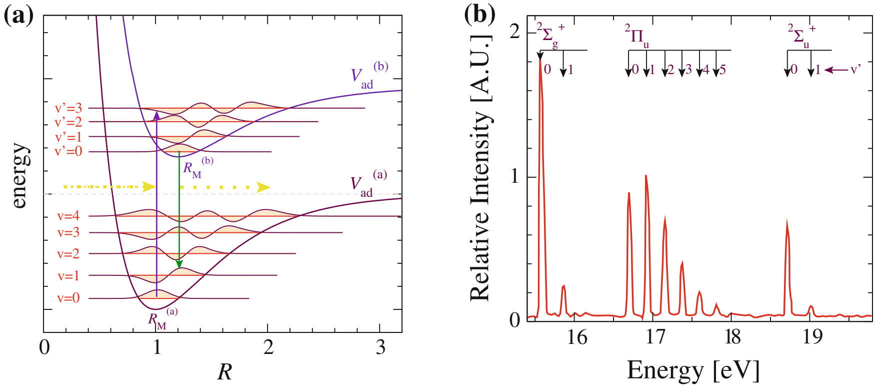


Fig. 3.14 **a** A sketch of the adiabatic potentials of two different molecular electronic states, with the respective vibrational energies, the ground, and a few excited vibrational wavefunctions. *Arrows* highlight the most intense vertical (Franck-Condon) transitions in absorption and in emission (fluorescence). **b** The observed photoemission spectrum of N_2, with the transition between the electronic ground state of N_2 and three different electronic states of N_2^+ labeled $^2\Sigma_g^+$, $^2\Pi_g$, $^2\Sigma_u^+$. This spectrum exhibits sequences of regularly-spaced vibrational satellites built on top of each electronic state of N_2^+ (Data from Ref. [27])

The intensities of the different satellites are distributed proportionally to the overlaps $|\langle v = 0|v'\rangle|^2$, involving the projection of the vibrational ground state $|v = 0\rangle$ of the initial adiabatic potential $V_{\text{ad}}^{(a)}$ on each of the final vibrational eigenstates $|v'\rangle$ of $V_{\text{ad}}^{(b)}$. The most probable transitions involve those $|v'\rangle$ states with large amplitude in the region of the initial minimum R'_M, as illustrated in Fig. 3.14a (*Franck-Condon principle*). Accordingly, the spectral intensity shared by a large number of vibrational satellites indicates a large displacement of the equilibrium position in the electronic transition. For example, the spectrum of Fig. 3.14b indicates that, in the electronic transition from the ground state of N_2 to N_2^+, R_M shifts more in going to the ${}^2\Pi_u$ state than to either of ${}^2\Sigma_g^+$ or ${}^2\Sigma_u^+$. Rotational structures accompanying the electronic-vibrational structures are complicated by the change in momenta of inertia and by occasional changes in electronic angular momentum.

3.3.3 Zero-Point Effects

We conclude this section with an intriguing detail: the bond energy of a diatom is slightly less than the depth $V_\mathrm{ad}(+\infty) - V_\mathrm{ad}(R_\mathrm{M})$ of the adiabatic potential well. The bond energy would coincide with the well depth $V_\mathrm{ad}(+\infty) - V_\mathrm{ad}(R_\mathrm{M})$ if the nuclear masses were infinite, or equivalently if the ions moved according to classical mechanics. Due to the quantum zero-point motion associated to Heisenberg's uncertainty principle (see Appendix B.2.1), the actual ground-state energy includes a vibrational contribution, that in the harmonic approximation equals $\mathscr{E}_\mathrm{vib}(0) = \hbar\omega/2$, see Eq. (3.21). As illustrated in Fig. 3.15, when zero-point energy is accounted for, the actual binding energy reduces to $E_b = V_\mathrm{ad}(+\infty) - [V_\mathrm{ad}(R_\mathrm{M}) + \mathscr{E}_\mathrm{vib}(0)]$. Experimentally, this usually small effect can be probed by changing the nuclear isotopic masses, thus modifying the vibrational frequency $\omega \propto \mu^{-1/2}$, without affecting $V_\mathrm{ad}(R)$.

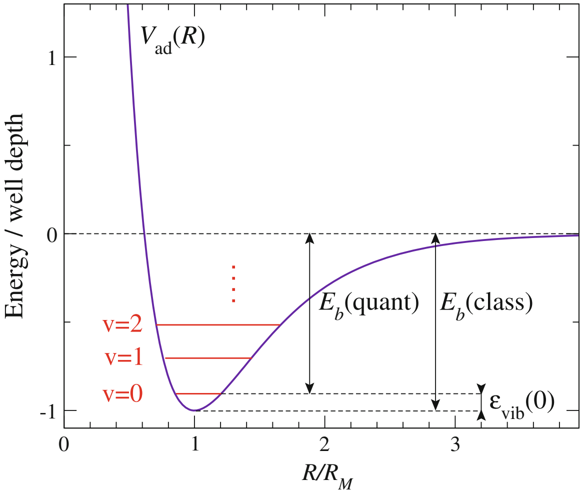


Fig. 3.15 The role of the zero-point vibrational motion in the precise definition of the bond energy of a diatomic molecule. The zero-point energy $\mathscr{E}_\mathrm{vib}(0)$ is a consequence of Heisenberg's uncertainty principle, namely the nonzero quantum kinetic term $\langle T_n \rangle$ in a position-localized state. The zero-point energy decreases as the atomic masses M_α increase: $\mathscr{E}_\mathrm{vib}(0)$ would eventually vanish in the limit of a classical adiabatic atomic motion

This zero-point effect is spectacular in ^{4}He$_2$. This dimer is so extremely weakly bound (the well depth is approximately 900 μeV, see Table 3.1) that a single bound $v = 0$ "vibrational" level exists [28], with a zero-point energy $\mathscr{E}_{\text{vib}}(0)$ that almost entirey balances the adiabatic attraction, to a total net binding energy $E_b \simeq 0.1$ μeV only! Of the 25 % lighter ^{3}He$_2$, no bound state is observed.

The present chapter summarizes a few basic ideas and experimental evidence in the field of molecular physics. The concepts of adiabatic separation and chemical bonding stand at the heart of all physics and chemistry of matter: they open the way to understanding the dynamics of systems composed of more than a single atom, namely molecules, as well as polymers, clusters, solids… The basic concepts sketched in these pages only scratch the surface of the broad field of molecular spectroscopy, which provides detailed quantitative information about the geometry and dynamics of diatomic and polyatomic molecules [6, 8, 9]. Beside spectroscopic characterization, molecules are synthesized transformed and investigated by means of all sorts of chemical reactions: these can mostly be analyzed conceptually in terms of the dynamics of atoms guided by suitable multi-dimensional potential energy surfaces $V_{\text{ad}}(R)$. The qualitative understanding and quantitative study of these phenomena constitutes the hard core of chemistry and chemical physics.

Problems

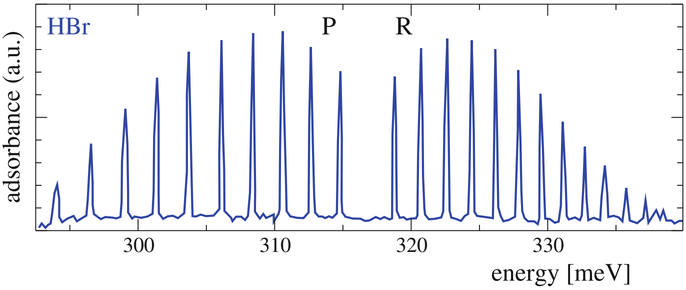
A $\star$ marks advanced problems. Problems labeled $^{\mathbf{S}}$ involve concepts of statistical physics from Chap. 4.

3.1 $^{\mathbf{S}}$ Consider a H^{80}Br gas sample at temperature 400 K. The equilibrium interatomic distance is $R_{\text{M}} = 141$ pm and the wave number of the vibrational transition $\bar{\nu} = 2{,}650\,\text{cm}^{-1}$. Evaluate:

- the rotational quantum number l of the most populated state;
- the wave number of the most intense transitions of the P and R branches of the absorption spectrum.

3.2 $^{\mathbf{S}}$ The vibrational wave number of the isotopically pure H^{79}Br molecule is $\bar{\nu} = 2{,}650\,\text{cm}^{-1}$. Evaluate the force needed to stretch the bond, displacing the two nuclei from their equilibrium distance R_{M} by 1 pm. Evaluate also the vibrational contribution to the specific heat capacity at temperature 350 K.

3.3 $^{\mathbf{S}}$ The figure below records an infrared absorption spectrum of gas-phase H^{80}Br.



Based on this experimental spectrum, determine:

- The elastic constant of the adiabatic potential in the harmonic approximation;
- the equilibrium distance between the H and Br nuclei;
- an estimate (with a 30 % precision) of the sample temperature.

3.4 $^{\mathrm{S}}$ Consider fluorine gas at temperature 5,000 K. Using the harmonic approximation, estimate the number of bound vibrational levels for a F_2 molecule, knowing that the vibrational transition is observed at $\bar{\nu} = 892\,\mathrm{cm}^{-1}$ and that the molecular binding energy is 1.5 eV. Write the vibrational partition function in the harmonic approximation, and estimate the fraction of bound molecules and the fraction of dissociated molecule. For the latter, assume simply that the dissociated molecules are represented by the harmonic vibrational levels above the dissociation threshold.

3.5 Evaluate the force (in Newton) necessary to pull the two nuclei of a N_2 molecule 1 pm apart, starting from the equilibrium separation $R_{\mathrm{M}} = 109\,\mathrm{pm}$. The vibrational frequency is observed at a wave number $2{,}360\,\mathrm{cm}^{-1}$. Evaluate also the excitation energy of the rotational state with angular momentum $l = 2$ relative to the ground state $l = 0$.

3.6 In the rovibrational spectrum of gas-phase HCl the transition $(v = 0, l = 2) \rightarrow (v = 1, l = 3)$ is observed at frequency 88,380 GHz. Evaluate the frequency of the homologous line in the DCl spectrum, knowing that the molecular equilibrium separation is 1.27 Å. Additionally, estimate the difference in dissociation energy between the two molecular species.

3.7 $^{\mathrm{S}}$ Evaluate the total molar specific heat capacity of a gas of CO molecules at temperature 500 K, knowing that the lowest pure rotational line is observed at frequency 115.271 GHz, and that the vibrational frequency is 64,100 GHz.

3.8 Model the adiabatic potential energy as a function of the distance R between a carbon and a oxygen nucleus with the Lennard-Jones potential:

$$V(R) = \varepsilon \left[\left(\frac{\sigma}{R} \right)^{12} - \left(\frac{\sigma}{R} \right)^{6} \right].$$

Evaluate the parameters ε and σ so as to reproduce the spectroscopic values of the vibrational quantum $\bar{\nu} = 2{,}162\,\mathrm{cm}^{-1}$, and the $3.8\,\mathrm{cm}^{-1}$ separation between the rotational lines of the $^{12}C^{16}O$ molecule.

3.9 $^{\mathrm{S}}$ The equilibrium interatomic distance of the HCl molecule is 127 pm, and the curvature of the adiabatic potential at this equilibrium distance amounts to 590 kg/s^2. For a gas-phase sample of ^{1}H^{37}Cl at equilibrium at temperature 390 K, determine the frequency in Hz of the most intense rovibrational absorption line in the P branch.

3.10 * Refining the standard rigid-rotor approximation, the actual equilibrium distance of a diatomic molecule does not coincide with the adiabatic minimum of $V_{\mathrm{ad}}(R)$, but with the minimum of $V_{\mathrm{ad}}(R) + V_{\mathrm{centrif}}(R)$, where the centrifugal contribution to the radial equation depends on the interatomic separation R and on the rotational quantum number l as well. As a consequence, the momentum of inertia depends on l, so that the purely rotational spectrum gets distorted compared to the rigid-rotor spectrum. Consider the ^{1}H^{127}I molecule and assume

$$V_{\mathrm{ad}}(R) = \varepsilon\left[\left(\frac{\sigma}{R}\right)^4 - 2\left(\frac{\sigma}{R}\right)^2\right],$$

with $\varepsilon = 3.0$ eV and $\sigma = 161$ pm. Evaluate the l value for which the corrected rotational energy $\mathscr{E}_{\mathrm{rot}}(l)$ differs by 1 % compared to the prediction of the rigid-rotor formula.

3.11 * Evaluate the mean distance between the two protons in the H$_2$ molecule and the mean square fluctuation of that distance due to the zero-point radial motion in the ground state, assuming the harmonic approximation. Also compute the ratio of these lengths. The vibrational wave number (4,395 cm^{-1}) and the rotational spectroscopic constant ($\hbar^2/2I = 61.0\,\mathrm{cm}^{-1}$) have been determined experimentally.
[Recall the expression $x_0 = [\hbar/(\omega\mu)]^{1/2}$ for the characteristic length scale of a harmonic oscillator of angular frequency ω and mass μ.]

3.12 * Consider the transition between two electronic states 1 and 2 of the diatomic molecule NH. Assume that the corresponding adiabatic potential energies $V_{\mathrm{ad}}^{(i)}(R)$ [for $i = 1, 2$] as a function of the interatomic distance R can be represented by the (Morse) form:

$$V_{\mathrm{ad}}^{(i)}(R) = E_{b\,i}\left[e^{-2a_i(R-R_{0\,i})} - 2e^{-a_i(R-R_{0\,i})}\right].$$

Assume the following values of the parameters: $E_{b1} = 5.9$ eV, $R_{01} = 95$ pm, $a_1 = 19.0\,(\mathrm{nm})^{-1}$, $E_{b2} = 3.2$ eV, $R_{02} = 107$ pm, $a_2 = 19.0\,(\mathrm{nm})^{-1}$. Adopting the harmonic approximation for the vibrational energies, estimate the wavelength of the radiation necessary to excite the transition from level 1 (ground vibrational state $v_1 = 0$) to level 2 (vibrational level $v_2 = 3$).

Chapter 4
Statistical Physics

The purpose of statistical physics is to relate average properties of "*macroscopic*" *objects* (thermodynamic quantities) to the fundamental interactions governing their *microscopic dynamics*. The previous chapters have discussed the way many detailed properties of individual atoms and molecules are accounted for based on the fundamental electromagnetic interactions driving the motion of the composing electrons and nuclei. The number of electrons and nuclei in atoms and small molecules does not exceed few hundred or, at most, few thousand. Macroscopic objects, as opposed to microscopic systems, are characterized by huge numbers of degrees of freedom. For example, a sodium-chloride crystal weighting 1 g, ready to be thrown in hot water for cooking *pasta*, is composed of about 3×10^{23} electrons plus approximately 10^{22} Na and 10^{22} Cl nuclei. One may as well conceive some (possibly approximate) wavefunction for the dynamics of such a huge number of degrees of freedom, but must also readily give up any hope to ever record the huge amount of information stored by the wavefunction describing the full detail of such a system. On the other hand, this limitation is not especially bad, since such intimate details of the dynamics of our salt crystal, because the individual motions of electrons and nuclei are probably boring and of little practical use. Physicists and materials scientists are more interested in the measurable *average macroscopic properties* of objects and substances, e.g. stiffness, tensile strength, heat capacity, heat and electrical conductivity, dielectric and magnetic susceptibilities, phase transitions…

4.1 Introductory Concepts

To draw a link between the microscopic dynamics and the macroscopic thermodynamic properties, equilibrium *statistical physics* borrows its mathematical tools from *the theory of probability* and from general *statistics*.

N. Manini, *Introduction to the Physics of Matter*,
Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-14382-8_4

4.1.1 Probability and Statistics

Statistics is rooted in the notion of probability. The naive notion of probability coincides with the relative number of observations. For example, after rolling a dice a large number N of times, we observe "two" N_2 times: the ratio N_2/N is an estimate of the probability P_2 of obtaining "two" in a single roll. However, we have also an a-priori idea of probability. We will assert that $P_2 = 1/6$, even against contradictory observation, unless we have evidence that the dice is loaded. Such a-priori notion of probability is an example of a rigorous definition of probability as a measure defined on a "space of events", such that the measure of all space equals unity. The basic assumptions are:

1. for non-intersecting sets of events A and B, the probability of $A \cup B$ (i.e. that any event in either A or B is realized) equals $P_A + P_B$;
2. given two spaces of events (which could intersect, or even coincide), two sets of events A and B belonging to the first and second space are called *independent* if the probability of (A and B) (i.e., an event occurs that satisfies the conditions for belonging to A and at the same time to B) is the product $P(A) \cdot P(B)$.

Both these properties are trivial when probability and relative number of observations are identified.

The two basic properties of probability sketched above allow us to relate the probabilities of complicated events to those of elementary events. For example, the probability of obtaining two ones when two dices are rolled independently is $1/6 \times 1/6 = 1/36$ (property 2); the probability for a four and a five is $1/36 + 1/36 = 1/18$, since both (a four on dice 1 and a five on dice 2) and (a five on dice 1 and a four on dice 2) are mutually exclusive events (property 1).

Statistics makes a wide use of *probability distributions*: these are lists of probabilities of mutually exclusive sets of events which cover the whole space of events. For example, when the two rolled dices are considered, the outcome may be grouped in equal numbers (6 mutually exclusive possibilities, each of probability $1/36$) or different numbers ($6 \times 5/2 = 15$ mutually exclusive possibilities, each of probability $1/18$): using property 1, this leads to a distribution $P_{\text{equal}} = 1/6$, and $P_{\text{different}} = 5/6$. Clearly, the events considered exhaust all space, and the sum of all the probabilities in the distribution equals unity. It is a useful exercise to work out the distribution $P_{\text{both even}}$, $P_{\text{even odd}}$, $P_{\text{both odd}}$ on the same space of events. Basically, a probability distribution makes probability a function of certain specifications of the events considered. Examples of popular distributions in statistical physics are the binomial distribution, the Poisson distribution, and the Gaussian distribution.

QM, even at the level of a single particle, or few of them, has an intrinsic statistical interpretation, in the probabilistic postulate of observation. When a measurement of observable A is done on a quantum system initially in some state $|i\rangle$, the system will be found in the eigenstate $|a\rangle$ of A associated to eigenvalue a, with probability $P_a(|i\rangle) = |\langle a|i\rangle|^2$. The quantum state $|i\rangle$ contains all probability distributions corresponding to all possible operators associated to conceivable measurements that could be carried out on the system. In fact, statistical physics is concerned only marginally with this statistical interpretation because, as long as the quantum-mechanical system is left undisturbed, the ket evolves from any given initial state according to the (deterministic) Schrödinger equation (1.7). On the contrary, the statistical description of a macroscopic system attempts to investigate its average properties without the need of a precise specification of the initial conditions, but rather assuming that all "reasonable" initial conditions could occur with equal probability. Ideally one would like to identify macroscopic properties with time averages of the microscopic observables along the evolution dictated by the internal dynamics.

Two basic assumptions reconcile the irrelevance of the initial conditions with the time-average viewpoint: equilibrium and ergodicity. *Equilibrium* requires that long ago the system has undergone some initial transient, and that now, at the time of interest, no systematic drift is occurring any more, all collective quantities (e.g. pressure) fluctuating in time around some well-defined average value. *Ergodicity* assumes that all kinds of states are randomly explored in a period of time short compared to the typical duration of measurements: subsequent times provide *independent random* realizations of an underlying probability distribution. A dishonest dice roller violates ergodicity by controlling accurately the initial conditions, rather than rolling blindly to generate successive truly random independent numbers. A similar violation may occasionally occur in statistical physics, as illustrated in the example of H_2 nuclear spin. H_2 molecules occur with total nuclear spin 1 (orthohydrogen) or 0 (parahydrogen). These states have essentially the same energy, thus they should all occur with equal probability: one expects to observe a ratio of ortho- to parahydrogen matching their degeneracy ratio 3:1. However ortho-para inter-conversion is rather difficult: ordinary molecular collisions conserve nuclear spins, thus leaving the total abundance of each species unaltered. Normally, inter-conversions occur at the walls, in the neighborhood of magnetic impurities. However, one could store the H_2 sample in a vessel where all magnetic impurities have been carefully removed. As a result, an anomalous abundance ratio of ortho- to parahydrogen can be stabilized for an extended time. Instead, if magnetic impurities are present, careless "shuffling" of the nuclear spins occurs, the expected 3:1 equilibrium distribution is readily recovered, and the system is ergodic. In brief, the ergodic hypothesis assumes that the system is sufficiently random that no conserved quantities prevents the access (in periods of time short compared to the duration of the experiment) to some major subset of the space of states.

When a system is at equilibrium and ergodic, time averages can safely be replaced by averages over a suitable probability distribution of the microscopic states. We now sketch the standard math used to describe a random distribution of quantum states.

4.1.2 Quantum Statistics and the Density Operator

We want to formalize our basic ignorance of the precise quantum state of a system, while describing correctly its statistical properties [17]. A system may be found in any of a number of quantum states $|a_1\rangle$, $|a_2\rangle$,... with probability w_1, w_2,... respectively. The states $|a_1\rangle$, $|a_2\rangle$,... need not be eigenstates of any observable, they need not be orthogonal, and in number they could even exceed the dimension of the Hilbert space of states. The normalization of probability requires that $\sum_i w_i = 1$.

On average, the measurement of an observable B on such a system should provide

$$[B] = \sum_i w_i \, \langle a_i|B|a_i\rangle = \sum_i w_i \sum_b b \, |\langle a_i|b\rangle|^2, \tag{4.1}$$

where the $\{|b\rangle\}$ is the basis of eigenkets of B, with eigenvalues b. We introduce the square brackets [ ] to indicate the ensemble average, i.e. the statistical mean of the quantum average values.

The averages (4.1) are naturally computed using the *statistical density operator*

$$\hat{\rho} = \sum_i w_i \, |a_i\rangle\langle a_i|, \tag{4.2}$$

as

$$\begin{aligned}[B] &= \sum_i w_i \langle a_i|B|a_i\rangle = \sum_i w_i \sum_b b\langle b|a_i\rangle\langle a_i|b\rangle = \sum_b b\langle b|\left(\sum_i w_i|a_i\rangle\langle a_i|\right)|b\rangle \\ &= \sum_b b\langle b|\hat{\rho}|b\rangle = \sum_b \langle b|\hat{\rho}B|b\rangle = \mathrm{Tr}(\hat{\rho}B).\end{aligned} \tag{4.3}$$

The density operator collects all dynamical and statistical properties of the system. $\hat{\rho}$ is self-adjoint: it can therefore be diagonalized. On its diagonal basis, the density operator is expressed as:

$$\hat{\rho} = \sum_m P_m \, |\rho_m\rangle\langle\rho_m|, \tag{4.4}$$

where the kets $|\rho_m\rangle$ form a complete orthonormal basis of the Hilbert space [in contrast to the $|a_i\rangle$ of the definition (4.2)], associated to positive eigenvalues P_m. Note the normalization

$$\begin{aligned}\mathrm{Tr}(\hat{\rho}) &= \sum_b \langle b|\hat{\rho}|b\rangle = \sum_{b\,i} w_i\langle b|a_i\rangle\langle a_i|b\rangle = \sum_i w_i \sum_b \langle a_i|b\rangle\langle b|a_i\rangle \\ &= \sum_i w_i\langle a_i|a_i\rangle = 1.\end{aligned} \tag{4.5}$$

indicating that also $\sum_m P_m = 1$: the eigenvalues P_m can be interpreted as probabilities.

The special case of a statistical operator describing a single *pure state* $\hat{\rho} = |a_1\rangle\langle a_1|$ falls back to standard deterministic QM: $[B] = \mathrm{Tr}(\hat{\rho}B) = \langle a_1|B|a_1\rangle = \langle B\rangle$. In this case (and only in this case) $\hat{\rho}$ has the property of a projector $\hat{\rho}^2 = \hat{\rho}$.

4.2 Equilibrium Ensembles

Consider a macroscopic isolated system (the "universe") U, of which the energy is known (*microcanonical ensemble*) up to an uncertainty ΔE. The fundamental *postulate of equilibrium statistical physics* follows from the ergodic hypothesis: *all*

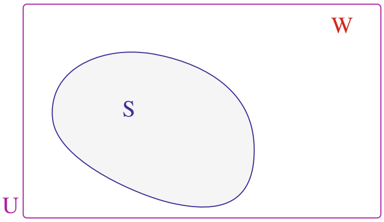


Fig. 4.1 The whole "universe" U partitioned into two weakly-interacting regions: a "system" S plus the "rest of the universe" W

quantum states of U are equally likely. In other words, the probability of any state $|i\rangle$ equals a constant $P_i^{\mathrm{U}} = 1/\Omega_{\mathrm{U}}(E;\,\Delta E)$ if the energy E_{U} of $|i\rangle$ is in the range $E \leq E_{\mathrm{U}} \leq E + \Delta E$, or vanishes otherwise. $\Omega_{\mathrm{U}}(E;\,\Delta E)$ equals the number of individual states of U in the selected energy range $E \leq E_{\mathrm{U}} \leq E+\Delta E$. In this way, the probability P_i^{U} is correctly normalized to unity. From this very "democratic" postulate, we derive the probability distribution of a system S in thermal equilibrium with the rest of the universe W (Fig. 4.1). The total Hamiltonian $H_{\mathrm{U}} = H_{\mathrm{S}} + H_{\mathrm{W}} + H_{\mathrm{SW}}$ is assumed to involve an extremely weak interaction term $H_{\mathrm{SW}} \approx 0$, so that $E_{\mathrm{U}} \approx E_{\mathrm{S}} + E_{\mathrm{W}}$.

We address the specific problem: What is the probability P_m of finding the system S in a given quantum state $|m\rangle$ of energy E_m? As the coupling between S and W is weak, we can assume that the states in S and in W are distributed *independently* at random. The only correlations between S and W are induced by the energy conservation. Accordingly,

$$P^{\mathrm{U}} = P_m^{\mathrm{S}}\, P^{\mathrm{W}}. \tag{4.6}$$

Microcanonical distributions govern both U and W individually: like U-states occur with probability $P^{\mathrm{U}} = 1/\Omega_{\mathrm{U}}(E;\,\Delta E)$, in W each state has probability $P^{\mathrm{W}} = 1/\Omega_{\mathrm{W}}(E - E_m;\,\Delta E)$. Here $E - E_m$ is the residual energy of W, once the energy E_m of S has been taken into account. From Eq. (4.6) we extract the probability distribution of S:

$$P_m^{\mathrm{S}} = \frac{P^{\mathrm{U}}}{P^{\mathrm{W}}} = \frac{1/\Omega_{\mathrm{U}}(E;\,\Delta E)}{1/\Omega_{\mathrm{W}}(E - E_m;\,\Delta E)} = \frac{\Omega_{\mathrm{W}}(E - E_m;\,\Delta E)}{\Omega_{\mathrm{U}}(E;\,\Delta E)}. \tag{4.7}$$

Observe that for most U-states the energy E_m in the system S is a tiny fraction of the total energy of the universe E: one can Taylor-expand the numerator (rather: its logarithm[1]) around E:

$$\ln \Omega_{\mathrm{W}}(E - E_m;\,\Delta E) = \ln \Omega_{\mathrm{W}}(E;\,\Delta E) - \beta E_m + \ldots, \tag{4.8}$$

---

[1] The number Ω_{W} of microstates of a huge "rest of the universe" increases roughly exponentially with its energy, thus the expansion of $\log \Omega_{\mathrm{W}}$ is far more accurate and better convergent than that of Ω_{W}.

where the slope $\beta = \partial \ln \Omega_{\mathrm{W}}(E'; \Delta E)/\partial E'\big|_{E'=E}$ has dimensions of inverse energy. The linear approximation for the logarithm is exceedingly good as long as $E_m \ll E_{\mathrm{W}}$. We substitute it back into Eq. (4.7) and obtain

$$P_m^{\mathrm{S}} = \frac{\Omega_{\mathrm{W}}(E; \Delta E)}{\Omega_{\mathrm{U}}(E; \Delta E)}\, e^{-\beta E_m}, \tag{4.9}$$

which is a very remarkable result for at least two reasons:

- the whole dependence on the state $|m\rangle$ of S is confined to the exponential factor, as the normalization factor $Z^{-1} = \Omega_{\mathrm{W}}(E; \Delta E)/\Omega_{\mathrm{U}}(E; \Delta E)$ is *independent of* $|m\rangle$;
- the *probability* of the state $|m\rangle$ of S *is a function of its energy* E_m *only*, through a simple exponential.

We rewrite Eq. (4.9) inserting the normalization factor Z^{-1} and omitting the label S:

$$P_m = P(E_m) = \frac{e^{-\beta E_m}}{Z}. \tag{4.10}$$

Equation (4.10) describes the *Boltzmann equilibrium probability distribution* for the states of a generic system in weak thermal contact with a huge environment. This distribution is associated to the Gibbs *canonical ensemble*.

As the probability distribution (4.10) is necessarily normalized, Z can alternatively be expressed entirely in terms of properties of the system S under study, namely

$$Z = \sum_m e^{-\beta E_m} = \mathrm{Tr}\left(e^{-\beta H}\right), \tag{4.11}$$

without any reference to U and W (thus, we write H in place of H_{S}). Z is usually called *partition function*. Note that the sum in Eq. (4.11) involves all microstates including, in particular, all same-energy components of degenerate levels. An equivalent formulation of the same sum is:

$$Z = \sum_E n_{\mathrm{deg}}(E) e^{-\beta E}. \tag{4.12}$$

where $n_{\mathrm{deg}}(E)$ is the number of degenerate states with energy E. Note also that the trace in Eq. (4.11) can be taken on any basis of convenience.

The density operator associated to the Boltzmann equilibrium probability distribution is *diagonal in the energy representation*:

$$\hat{\rho}_{\mathrm{eq}} = \sum_m \frac{e^{-\beta E_m}}{Z}\, |m\rangle\langle m| = \frac{1}{Z}\, e^{-\beta H}. \tag{4.13}$$

This same density operator can however be written on any basis, by applying a suitable unitary transformation.

4.2.1 Connection to Thermodynamics

Assume that the system S can be partitioned into two weakly interacting subsystems S_1 and S_2 ($E_{m_1 m_2} \approx E_{1m_1} + E_{2m_2}$). When thermal equilibrium is established between S_1, S_2, and the rest of the world W, the distribution

$$P_{m_1 m_2} = Z^{-1} e^{-\beta E_{m_1 m_2}} = \frac{e^{-\beta(E_{1m_1}+E_{2m_2})}}{\sum_{n_1 n_2} e^{-\beta(E_{1n_1}+E_{2n_2})}},$$

can be factorized into

$$P_{m_1 m_2} = \frac{e^{-\beta E_{1m_1}}}{\sum_{n_1} e^{-\beta E_{1n_1}}} \cdot \frac{e^{-\beta E_{2m_2}}}{\sum_{n_2} e^{-\beta E_{2n_2}}} = \frac{e^{-\beta E_{1m_1}}}{Z_1} \cdot \frac{e^{-\beta E_{2m_2}}}{Z_2}. \tag{4.14}$$

This decomposition shows that at equilibrium S_1 and S_2 follow independent Boltzmann distributions $P^{S_i}_{m_i} = e^{-\beta E_{im_i}}/Z_i$, with the *same* β parameter. This observation suggests that the intensive quantity β might be a function of temperature T.

The natural statistical definition for the *internal energy* U of thermodynamics is the average of the energy operator

$$U = [H] = \mathrm{Tr}(\hat{\rho} H) = \sum_m E_m P_m = \frac{1}{Z} \sum_m E_m e^{-\beta E_m}. \tag{4.15}$$

Deriving U with respect to β yields minus the squared energy fluctuation:

$$\begin{aligned} \frac{\partial U}{\partial \beta} &= \frac{-Z \sum_m E_m^2 e^{-\beta E_m} + \left(\sum_m E_m e^{-\beta E_m}\right)^2}{Z^2} \\ &= -\left([H^2] - [H]^2\right) = -\left[\,(H - [H])^2\,\right]. \end{aligned} \tag{4.16}$$

Thus, $\partial U/\partial \beta \leq 0$, i.e. the internal energy decreases when β increases. This observation suggests that β and temperature might be inversely related. The determination of the precise relationship is sketched below.

Firstly, note that Z is a multiplicative function [$Z = Z_1 Z_2$ for a system composed of two subsystems, see Eq. (4.14)], thus its logarithm is an *additive function* ($\ln Z = \ln Z_1 + \ln Z_2$). $\ln Z$ must then represent an *extensive* thermodynamic quantity. Define the function

$$F = -\frac{\ln Z}{\beta}. \tag{4.17}$$

Using this definition, we derive a relation between the β-derivative of the extensive quantity $\ln Z$ and the internal energy U:

$$\begin{aligned} \left.\frac{\partial(\beta F)}{\partial \beta}\right|_{V,N} &= -\frac{\partial(\ln Z)}{\partial \beta} = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{\frac{\partial}{\partial \beta}\sum_m e^{-\beta E_m}}{Z} \\ &= -\frac{\sum_m \frac{\partial}{\partial \beta} e^{-\beta E_m}}{Z} = \frac{\sum_m E_m e^{-\beta E_m}}{Z} = [H] = U, \end{aligned} \tag{4.18}$$

namely the average energy, as defined in Eq. (4.15). Equation (4.18) identifies the derivative of βF w.r.t. β with the internal energy. This identification recalls a known thermodynamic identity involving the free energy and temperature: starting from the basic definition of the free energy $F = U - TS$ (where S is entropy), one finds

$$U = F + TS = F - T\,\frac{\partial F}{\partial T} = -T^2\left(-T^{-2}F + T^{-1}\frac{\partial F}{\partial T}\right) = -T^2\,\frac{\partial (T^{-1}F)}{\partial T}. \tag{4.19}$$

This expression can be written more compactly in terms of a derivative w.r.t. the inverse temperature:

$$\frac{\partial (T^{-1}F)}{\partial (T^{-1})} = U. \tag{4.20}$$

In Eqs. (4.19) and (4.20) all T-derivatives are carried out at constant number of particles N and volume V. By comparing Eqs. (4.18) and (4.20) we see that

- it is perfectly natural to identify the extensive quantity F defined statistically in Eq. (4.17) with the free energy F of thermodynamics;
- β must then be proportional to the inverse temperature T^{-1}.

The proportionality constant between T and β^{-1} is known as the Boltzmann constant k_B, and represents the numerical conversion factor between temperature [K] and energy [J]:

$$\beta = \frac{1}{k_\mathrm{B}T}. \tag{4.21}$$

The comparison of Eq. (4.47) with the empirical definition of temperature through the ideal-gas thermometer provides the relation $k_\mathrm{B} = R/N_\mathrm{A}$ with the gas constant R and the Avogadro constant N_A. The resulting $k_\mathrm{B} = 1.38065 \times 10^{-23}$ J/K $=$ 86.1733 μeV/K. In statistical physics, temperature usually appears in the energy combination $k_\mathrm{B}T$ or in the β notation.

The physical meaning of temperature is now clear: *the probability distribution of the microstates of any system in thermal equilibrium is determined* uniquely *by the energies of these states*, in a straightforward way: at a given temperature T, *the probability $P(E_m)$ of a given state of energy E_m is proportional to* $\exp(-\beta E_m)$. Thus, $-\beta = -(k_\mathrm{B}T)^{-1}$ is the slope of the straight line representing $P(E_m)$ as a function of E_m in a lin-log plot, as in Fig. 4.2.

Note that, due to the normalization factor Z^{-1} in Eq. (4.10), the energies of all states in the system determine the equilibrium probability P_m of the occurrence of a given energy eigenstate $|m\rangle$. However, only the energy *difference* between two states $|m\rangle$ and $|n\rangle$ determines their probability *ratio* $P_m/P_n = \exp[\beta(E_n - E_m)]$.

The discussed relations, in particular Eqs. (4.17) and (4.21), establish an explicit link between statistical physics and thermodynamics, i.e. between the microscopic dynamics and macroscopic observable average properties. Based on Eq. (4.17), one

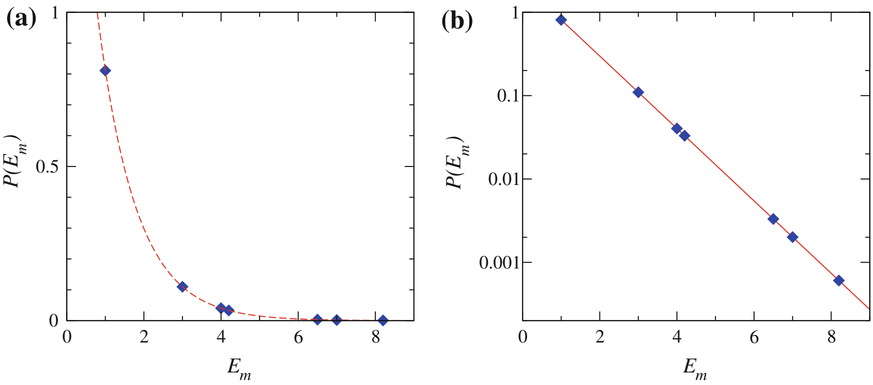


Fig. 4.2 The statistical meaning of temperature: **a** in a system at equilibrium with the "rest of the universe" the probability $P(E_m)$ of individual microstates $|m\rangle$ decays exponentially with their energy E_m. **b** The slope of the *straight line* representing $P(E_m)$ as a function of energy in a lin-log scale is precisely $-\beta = -(k_B T)^{-1}$

can link all thermodynamic properties of a system at equilibrium purely to its partition function Z. We list here the expressions for several fundamental extensive and intensive thermodynamic quantities:

$$\text{free energy } F = -k_B T \ \ln Z = -\frac{\ln Z}{\beta} \tag{4.22}$$

$$\text{internal energy} \quad U = F + TS = \left.\frac{\partial(\beta F)}{\partial \beta}\right|_{V,N} = -\left.\frac{\partial(\ln Z)}{\partial \beta}\right|_{V,N} \tag{4.23}$$

$$\text{entropy} \quad S = -\left.\frac{\partial F}{\partial T}\right|_{V,N} = k_B T \frac{\partial \ln Z}{\partial T} + k_B \ln Z = \frac{U}{T} + k_B \ln Z \tag{4.24}$$

$$\text{heat capacity} \quad C_V = \left.\frac{\partial U}{\partial T}\right|_{V,N} = -k_B\, \beta^2 \left.\frac{\partial U}{\partial \beta}\right|_{V,N} = k_B\, \beta^2 \frac{\partial^2 \ln Z}{\partial \beta^2} \tag{4.25}$$

$$\text{pressure} \quad P = -\left.\frac{\partial F}{\partial V}\right|_{\beta,N} = \left(\frac{N}{V}\right)^2 \left.\frac{\partial (F/N)}{\partial (N/V)}\right|_{\beta} \tag{4.26}$$

$$\text{free enthalpy} \quad G = F + PV \tag{4.27}$$

$$\text{enthalpy} \quad H = F + PV + TS. \tag{4.28}$$

4.2.2 Entropy and the Second Principle

The connection (4.24) of entropy and statistics applies at equilibrium only. An entirely *general definition of entropy* is available, which applies for any statistical distribution of the quantum states defined by an arbitrary density operator $\hat{\rho}$, whether it represents an equilibrium ensemble or not:

$$S = -k_{\mathrm{B}} \mathrm{Tr}(\hat{\rho} \ln \hat{\rho}). \tag{4.29}$$

This definition conforms to the intuitive idea of entropy as a measure of disorder. Consider the basis $|\rho_m\rangle$ where $\hat{\rho}$ is diagonal. Here, $\hat{\rho} = \sum_m P_m |\rho_m\rangle\langle\rho_m|$. In this basis

$$S = -k_{\mathrm{B}} \sum_m P_m \ln P_m. \tag{4.30}$$

We can now compute S for the most ordered distribution, namely a pure state (i.e. $P_m = 1$ for a single state m and 0 for all others), and obtain $S = 0$ since all terms vanish. S increases when several states have nonzero probabilities. For the extreme limit of a completely random distribution with probability $P_m = \Omega^{-1}$ for a number Ω of equally-likely states, we obtain $S = -k_{\mathrm{B}} \sum_m^{\Omega} \left[\Omega^{-1} \ln(\Omega^{-1})\right] = k_{\mathrm{B}} \ln(\Omega)$. We conclude that *entropy is a logarithmic measure of the number of states* that the system accesses.

When the system is at equilibrium, the general statistical definition (4.29) coincides with Eq. (4.24). This is readily verified by substituting the equilibrium density operator $\hat{\rho}_{\mathrm{eq}}$ of Eq. (4.13) into (4.29), i.e. Eq. (4.10) into (4.30):

$$\begin{aligned} S &= -k_{\mathrm{B}} \sum_m P_m \ln P_m = -k_{\mathrm{B}} \sum_m \frac{e^{-\beta E_m}}{Z} \ln \frac{e^{-\beta E_m}}{Z} \\ &= -k_{\mathrm{B}} \frac{\sum_m e^{-\beta E_m} (\ln e^{-\beta E_m} - \ln Z)}{Z} = k_{\mathrm{B}} \frac{\sum_m e^{-\beta E_m} (\beta E_m + \ln Z)}{Z} \\ &= k_{\mathrm{B}} \beta \, [H] + k_{\mathrm{B}} \frac{Z \ln Z}{Z} = \frac{U}{T} + k_{\mathrm{B}} \ln Z. \end{aligned} \tag{4.31}$$

Moreover, it is possible to show [29] that *if* $\hat{\rho}$ *coincides with the* Gibbs-Boltzmann *equilibrium density matrix* $\hat{\rho}_{\mathrm{eq}}$, Eq. (4.13), the system *entropy*, defined in Eq. (4.29), *is maximum* under the constraint of assigned internal energy $\mathrm{Tr}(\hat{\rho} H) = U$, volume V, and number of particles N. Any generic density operator $\hat{\rho}_{\mathrm{gen}}$, respecting the constraints on U, V, and N, yields an entropy S_{gen} not exceeding the S_{eq} of the equilibrium distribution $\hat{\rho}_{\mathrm{eq}}$, Eq. (4.13). Any isolated system is observed to evolve spontaneously toward equilibrium: accordingly, this maximum-entropy result shows the consistency of statistical physics with the *second principle of thermodynamics*. In a spontaneous transformation toward equilibrium $\hat{\rho}_{\mathrm{gen}} \to \hat{\rho}_{\mathrm{eq}}$, the entropy increases from S_{gen} up to S_{eq}. This result can be extended to conditions where U, V, or N vary (not-isolated systems), but for our purposes it suffices to retain that the experimental fact of entropy increase in the approach to equilibrium is rooted in statistics.

For an in-depth analysis of the concepts sketched in the present section, we refer to specialized textbooks, e.g. Refs. [17, 29, 30].

4.3 Ideal Systems

Before attempting any understanding of the complicated macroscopic properties of structured objects such as a bicycle or a bowl of soup, scientists have wisely chosen to investigate simpler systems: (usually) macroscopically homogeneous pure (or controllably "doped") bunches of atoms or molecules all of the same kind, or of few kinds. These simpler systems constitute the wide class of *materials*. The rationale behind studying materials is that the properties of a complex structured object can be understood in terms of the functionality of the individual pieces it is composed of, whose function, in turn, depends on their shapes and material properties. For example, the total heat capacity of a bicycle is practically identical to the sum of the heat capacities of the composing parts.

The macroscopic properties of a material are often studied in the limit of an infinitely broad sample, for which the interactions with the surrounding environment (e.g. the containing vessel) are sufficiently weak to justify the assumptions of Sect. 4.2. As the surface atoms/molecules directly interacting with the environment usually involve a layer about $\sim$1 nm thick, for any sample whose linear dimensions (all three of them) are larger than, say, $\sim$1 μm, the error induced by this *bulk approximation* should not be too bad.

Even with the idealization of a homogeneous bulk sample, the recipe

1. compute the spectrum of energies E_m and eigenstates $|m\rangle$ of the system;
2. at given temperature T [or $\beta = 1/(k_\mathrm{B}T)$], compute the partition function Z, Eq. (4.11);
3. generate the equilibrium density operator $\hat{\rho}$ using Eq. (4.13);
4. compute macroscopic average quantities as $[B] = \mathrm{Tr}(\hat{\rho}B)$, as in Eq. (4.3);

is not really applicable for any realistic system, due to the difficulty of step 1.

Luckily, for a few *ideal systems* the programme of statistical physics can be carried to satisfactory conclusion. By ideal systems we mean systems composed of individual components (particles) whose mutual interactions are negligible. Ideal systems have

$$H = \sum_{i=1}^{N} H_i, \tag{4.32}$$

where H_i governs the dynamics of a small set of degrees of freedom, e.g. the position and spin of one particle. The simplicity of ideal systems allows us to obtain exact partition functions, and thus to predict their thermodynamic properties by means of statistical methods.

No ideal system exists in nature (although photons and neutrinos are very good approximations to non-interacting particles). If one did exist, strictly speaking the methods of equilibrium statistical physics would be irrelevant for that system since it would not be ergodic, and could never reach equilibrium. However, several properties of *weakly-interacting* systems resemble those of ideal systems. Most properties of gaseous materials can be understood qualitatively, and often even quantitatively, in terms of properties of ideal systems. The reader should however be cautioned of the limited applicability and the risk of artifacts due to the idealized nature of non-interacting systems: a quantitative understanding of the thermodynamic properties of real materials usually requires more sophisticated methods, to account for interactions.

Consider an ideal system composed of a single type of identical noninteracting particles, e.g., electrons, or atoms, or molecules... The natural basis of states is a factored basis, Eqs. (2.60) and (2.61), where the α_i quantum numbers identify the state $|\alpha_i\rangle$ of particle i. Ignoring temporarily the permutational symmetry requirements (Sects. 2.2.1 and 2.2.2), the partition function

$$\begin{aligned} Z &\overset{?}{=} \sum_{\alpha_1\,\alpha_2\cdots\alpha_N} \exp[-\beta(\mathcal{E}_{\alpha_1} + \mathcal{E}_{\alpha_2} + \cdots + \mathcal{E}_{\alpha_N})] \\ &= \sum_{\alpha_1\,\alpha_2\cdots\alpha_N} e^{-\beta\mathcal{E}_{\alpha_1}} e^{-\beta\mathcal{E}_{\alpha_2}} \ldots e^{-\beta\mathcal{E}_{\alpha_N}}\,. \end{aligned} \tag{4.33}$$

(with $\mathcal{E}_\alpha$ indicating the single-particle eigenenergies of state $|\alpha\rangle$) factorizes into N single-particle terms. However, due to (anti)symmetrization, the exchange of any two quantum numbers α_i and α_j leads to the same state (possibly up to a sign) for the N identical particles. For fermions, with all different α_i's, the number of permutations of the quantum numbers giving the same state is $N!$. For bosons, where some α_i's can coincide, the number of permutations of the quantum numbers is smaller. To determine this number of permutations, first sort the single-particle states $|\alpha\rangle$, e.g. for increasing energy, say $|0\rangle$, $|1\rangle$, $|2\rangle$, ... Then count the numbers $n_0, n_1, n_2, \ldots$ of times each of these single-particle states appears in any given term in the sum (4.33). Clearly, the sum of these *occupation numbers* $n_0 + n_1 + n_2 + \cdots = N$. The number of permutations of a set of indexes $\alpha_1, \alpha_2, \ldots, \alpha_N$ is $N!/(n_0!\ n_1!\ n_2!\ \ldots)$. This expression holds for fermions as well, because all $n_\alpha! = 0! = 1$ or $1! = 1$.

We can now correct the partition function to remove multiple counts of the same states:

$$Z = \sum_{\alpha_1\,\alpha_2\cdots\alpha_N} \frac{n_0!\,n_1!\,n_2!\ldots}{N!} \exp\left(-\beta\sum_{i=1}^{N}\mathcal{E}_{\alpha_i}\right). \tag{4.34}$$

The sum of Eq. (4.34) extends over all possible α_i's for bosons, and over different α_i's respecting Pauli's principle in the case of fermions.

4.3.1 The High-Temperature Limit

At low temperature, the Boltzmann exponential factor tends to privilege those states with a total energy $\sum_{i=1}^{N}\mathcal{E}_{\alpha_i} = \sum_\alpha n_\alpha\mathcal{E}_\alpha$ as small as possible: relatively few

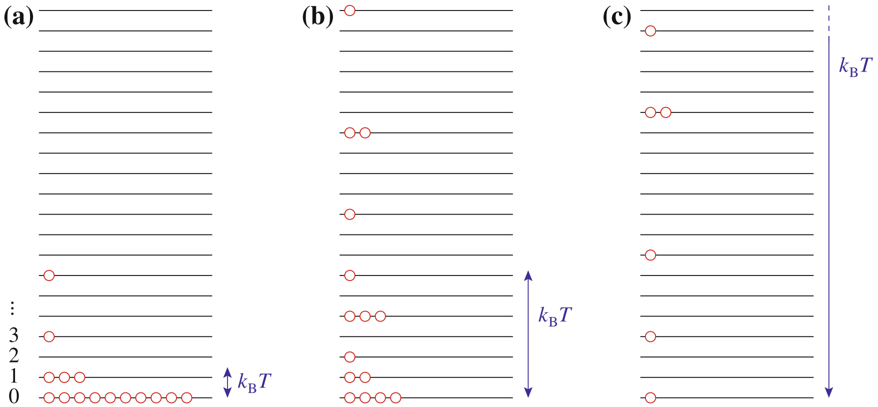


Fig. 4.3 Typical boson occupancies of single-particle energy eigenstates in a spectrum with no *upper* bound. **a** Low temperature. **b** Intermediate temperature. **c** High temperature, with most levels left unoccupied

low-energy single-particle states are significantly occupied, see Fig. 4.3a. We shall come back to this low-temperature regime in Sect. 4.3.2.

In contrast, at high temperature (small β), assuming, as is usually the case, that the single-particle spectrum has no upper bound, the occupation probabilities of the huge number ($\gg N$) of single-particle states with energy $\lesssim k_BT$ are very similar. Thus the probability that one individual state is occupied is very small (Fig. 4.3c). N-particle states with all different quantum numbers are overwhelmingly more numerous than those with two or more of them equal. As a result, at high temperature the main contribution to the boson partition function comes from states which have $n_\alpha = 0$ or 1 at most. All terms with $n_\alpha > 1$ in the sum of Eq. (4.34) contribute negligibly to a huge Z. Thus in Eq. (4.34) we approximate $n_0!\,n_1!\,n_2!\cdots \simeq 1$ and even in the fermionic case we include the unphysical terms with $n_\alpha > 1$ in the sum. Thus we obtain an approximate high-temperature partition function for bosons and fermions alike:

$$Z \simeq \frac{1}{N!} \sum_{\alpha_1\,\alpha_2\,\ldots\,\alpha_N}^{\text{unrestricted}} \exp\left(-\beta \sum_{i=1}^{N} \mathscr{E}_{\alpha_i}\right) = \frac{1}{N!}\, Z_1 \cdot Z_2 \cdot \cdots \cdot Z_N = \frac{(Z_1)^N}{N!}. \tag{4.35}$$

Here the single-particle partition function $Z_1 = \sum_\alpha \exp(-\beta \mathscr{E}_\alpha)$ sums the Boltzmann weights $\exp(-\beta \mathscr{E}_\alpha)$ for all one-particle states $|\alpha\rangle$. Z_1 should not be confused with the full Z of Eq. (4.11) which involves entire-system states $|m\rangle$ instead.

In this high-temperature limit, the free energy is

$$F = -\frac{\ln Z}{\beta} \simeq -k_BT \ln \frac{(Z_1)^N}{N!} \simeq -k_BT\left(N \ln Z_1 - N \ln \frac{N}{e}\right) = -Nk_BT \ln \frac{e\,Z_1}{N}, \tag{4.36}$$

where we made use of the Stirling approximation $\ln(N!) \simeq N \ln(N/e)$, which is very accurate for large N.

At so high a temperature that interactions among the N identical particles can be neglected, the system is usually in its gas phase. The center of mass of each particle translates freely. This translational degree of freedom can be separated from the internal ones, as sketched in Sect. 2.1 for the one-electron atom. In the single-particle Hamiltonian $H_{1\,\mathrm{tr}} + H_{1\,\mathrm{int}}$ the two terms commute. Accordingly, the single-particle partition function Z_1 can be factored

$$Z_1 = Z_{1\,\mathrm{tr}}\, Z_{1\,\mathrm{int}} \tag{4.37}$$

into a translational times an internal part. The latter describes the statistical dynamics of the internal degrees of freedom of the particle (including, e.g., its spin, and/or its rotations), and depends therefore on the specific spectrum of its excitations. In contrast, the *translational part is universal*: it depends uniquely on the particle mass M and the containing volume V.

4.3.1.1 Translational Degrees of Freedom

To obtain $Z_{1\,\mathrm{tr}}$, recall the spectrum of a freely translating particle contained in a macroscopically large cubic box of volume $V = L\times L\times L$. Impose periodic boundary conditions, i.e. the eigenfunction is the same at opposite faces of the box. The results [Eq. (4.42) onward] would not change if we assumed that the wavefunction vanishes at the boundary instead. The allowed values of the $u = x, y, z$ momentum components

$$p_u = \hbar k_u = \hbar \frac{2\pi}{L} n_u \qquad n_u = 0, \pm 1 \pm 2 \pm 3, \ldots \tag{4.38}$$

are associated to plane-wave eigenfunctions $\psi_{\mathbf{k}}(\mathbf{r}) = L^{-3/2} \exp(i\,\mathbf{k}\cdot\mathbf{r})$, of translational kinetic energy

$$\mathscr{E}_{\mathbf{n}} = \frac{|\mathbf{p}|^2}{2M} = \frac{\hbar^2|\mathbf{k}|^2}{2M} = \frac{(2\pi\hbar)^2\,(n_x^2+n_y^2+n_z^2)}{2ML^2}. \tag{4.39}$$

For macroscopically large L, the translational energy levels form a "continuum": the triple sum in the translational single-particle partition function

$$Z_{1\,\mathrm{tr}} = \sum_{\mathbf{n}} \exp(-\beta\mathscr{E}_{\mathbf{n}}) = \sum_{n_x\, n_y\, n_z} \exp\left(-\beta\frac{[2\pi\hbar]^2[n_x^2+n_y^2+n_z^2]}{2ML^2}\right) \tag{4.40}$$

is conveniently approximated by the integral

$$Z_{1\,\mathrm{tr}} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\frac{\beta[2\pi\hbar]^2}{2ML^2}[n_x^2+n_y^2+n_z^2]\right) dn_x\, dn_y\, dn_z. \tag{4.41}$$

This expression factorizes into the product of three identical Gaussian integrals.[2] In total,

$$Z_{1\,\mathrm{tr}} = \left(\frac{L}{\hbar} \sqrt{\frac{Mk_\mathrm{B}T}{2\pi}} \right)^3 = \left(\frac{L}{\Lambda} \right)^3 = \frac{V}{\Lambda^3}, \tag{4.42}$$

where we have introduced the *thermal length*

$$\Lambda = \hbar \sqrt{\frac{2\pi}{Mk_\mathrm{B}T}}. \tag{4.43}$$

We can now substitute the result (4.42) for the single-particle partition function into the global partition function (4.35) of the gas

$$Z = \frac{(Z_{1\,\mathrm{tr}})^N}{N!} \left(Z_{1\,\mathrm{int}}\right)^N = \frac{V^N}{N!\,\Lambda^{3N}} \left(Z_{1\,\mathrm{int}}\right)^N . \tag{4.44}$$

Derivation yields the translational contribution to the internal energy

$$\begin{aligned} U_\mathrm{tr} &= -\frac{\partial}{\partial\beta} \ln Z_\mathrm{tr} = -\frac{\partial}{\partial\beta} \ln \frac{V^N}{N!\,\Lambda^{3N}} = 3N \frac{\partial \ln \Lambda}{\partial\beta} \\ &= 3N \frac{\partial \ln \beta^{1/2}}{\partial\beta} = \frac{3}{2}\frac{N}{\beta} = \frac{3}{2} Nk_\mathrm{B}T. \end{aligned} \tag{4.45}$$

This result is equivalent to the experimentally well-established translational contribution $\frac{3}{2}k_\mathrm{B}$ to the heat capacity per molecule of high-temperature gases. Likewise, we obtain an expression for the free-energy contribution of the translational motion:

$$F = -N\,k_\mathrm{B}T\,\ln \frac{e\,Z_{1\,\mathrm{tr}} Z_{1\,\mathrm{int}}}{N} = -N\,k_\mathrm{B}T \left(\ln \frac{e\,V}{N\Lambda^3} + \ln Z_{1\,\mathrm{int}} \right). \tag{4.46}$$

$Z_{1\,\mathrm{int}}$ describes internal degrees of freedom: thus it can depend on T, but surely not on the volume V of the containing vessel. This observation and the definition of pressure (4.26) allow us to derive a remarkably general equation of state for the ideal gas:

$$\begin{aligned} P &= -\left.\frac{\partial F}{\partial V}\right|_{\beta,N} = -\frac{\partial}{\partial V} \left[-Nk_\mathrm{B}T \left(\ln \frac{e\,V}{N\Lambda^3} + \ln Z_{1\,\mathrm{int}} \right) \right]_{T,N} \\ &= Nk_\mathrm{B}T \frac{\partial}{\partial V} \ln \frac{e\,V}{N\Lambda^3}, \end{aligned}$$

---

[2] Recall that $\int_{-\infty}^{\infty} \exp\left(-x^2/[2a^2]\right) dx = \sqrt{2\pi}\,a$.

i.e.

$$P = \frac{Nk_\mathrm{B}T}{V}. \tag{4.47}$$

This relation, obtained on purely statistical grounds, is equivalent to the well-established equation of state of perfect gases $PV = nRT$, universally and quantitatively valid for atomic and molecular gases at high temperature and low density. The numerical value $k_\mathrm{B} = R/N_\mathrm{A}$ is determined by identifying $Nk_\mathrm{B}T = nRT$, given that n moles of gas contain nN_A particles.

Other quantities are accessible experimentally. For example, the single-particle *kinetic-energy distribution* for center-mass translations, yields the probability to find a particle in a given energy interval. To obtain this distribution, we need the energy density of translational states

$$g_\mathrm{tr}(\mathscr{E}) = \frac{M^{3/2}\, V}{\sqrt{2}\,\pi^2\,\hbar^3}\,\mathscr{E}^{1/2}, \tag{4.48}$$

as can be derived from the kinetic-energy expression (4.39).[3] Then, in the spirit of Eq. (4.12), the single-particle kinetic-energy probability distribution is the product of the density of states times the Boltzmann probability that a given state is occupied:

$$\frac{dP(\mathscr{E})}{d\mathscr{E}} = g_\mathrm{tr}(\mathscr{E})\,\frac{e^{-\beta\mathscr{E}}}{Z_{1\,\mathrm{tr}}} = \frac{M^{3/2}V}{\sqrt{2}\pi^2\hbar^3}\mathscr{E}^{1/2}\,\frac{\Lambda^3}{V}e^{-\beta\mathscr{E}} = \frac{2}{\sqrt{\pi}}\,\beta^{3/2}\,\mathscr{E}^{1/2}\,e^{-\beta\mathscr{E}}. \tag{4.49}$$

This probability distribution, drawn in Fig. 4.4, is remarkably universal: it does not even depend on the mass of the particles, but uniquely on temperature.

Similarly, we derive the translational *velocity distribution*. Every component of $\mathbf{v} = \mathbf{p}/M$ is Gaussian-distributed as

$$\frac{dP(v_u)}{dv_u} = \sqrt{\frac{\beta M}{2\pi}}\,\exp\left(-\beta\frac{Mv_u^2}{2}\right). \tag{4.50}$$

To obtain the distribution of speed $v = |\mathbf{v}|$, one can simply observe that the kinetic energy $\mathscr{E}$ and speed v are connected by $\mathscr{E} = \frac{M}{2}v^2$, and use the distribution (4.49):

---

[3] According to Eq. (4.39), the kinetic energy is proportional to the squared length of the $\mathbf{n}$ vector $\mathscr{E}_\mathbf{n} = A|\mathbf{n}|^2$. As the $\mathbf{n}$ points are evenly distributed with unit density, the number of states with energy at most $\mathscr{E}$ equals the volume of the sphere of radius $|\mathbf{n}| = \sqrt{\mathscr{E}/A}$ in $\mathbf{n}$-space. The density of states is the derivative of this number of states with respect to energy: $g_\mathrm{tr}(\mathscr{E}) = \frac{d}{d\mathscr{E}}\frac{4\pi}{3}\,(\mathscr{E}/A)^{3/2} = 2\pi\mathscr{E}^{1/2}/A^{3/2}$. This expression yields Eq. (4.48) after substituting the value of $A = (2\pi\hbar)^2/(2ML^2)$ [from Eq. (4.39)], and $V = L^3$.

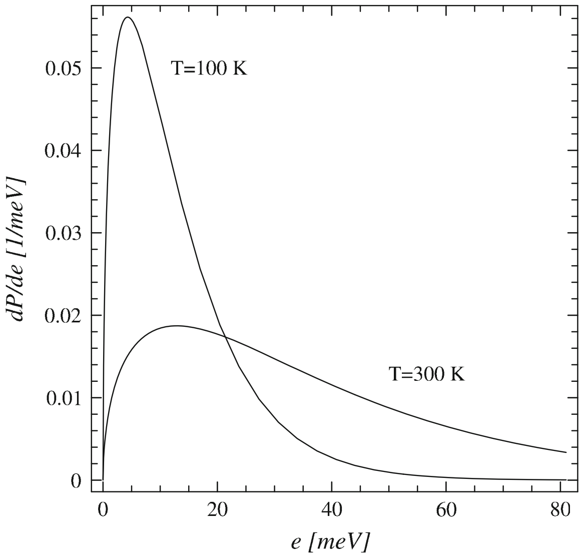


Fig. 4.4 The distribution—Eq. (4.49)—of the translational kinetic energy of the molecules in an ideal gas at a temperature high enough that Bose/Fermi statistics effects can be neglected

$$\begin{aligned}\frac{dP(v)}{dv} &= \frac{dP(\mathcal{E})}{d\mathcal{E}}\,\frac{d\mathcal{E}}{dv} = Mv\,\frac{dP(\mathcal{E})}{d\mathcal{E}} \\ &= Mv\frac{2}{\sqrt{\pi}}\beta^{3/2}\left(\frac{Mv^2}{2}\right)^{1/2} e^{-\beta\frac{M}{2}v^2} = \sqrt{\frac{2}{\pi}}\,(M\beta)^{3/2}\,v^2\,e^{-\frac{\beta M}{2}v^2}.\end{aligned} \tag{4.51}$$

Equation (4.51), the *Maxwell-Boltzmann* equilibrium speed distribution, is drawn in Fig. 4.5. The v^2 factor, owing the same "polar" origin discussed for the radial

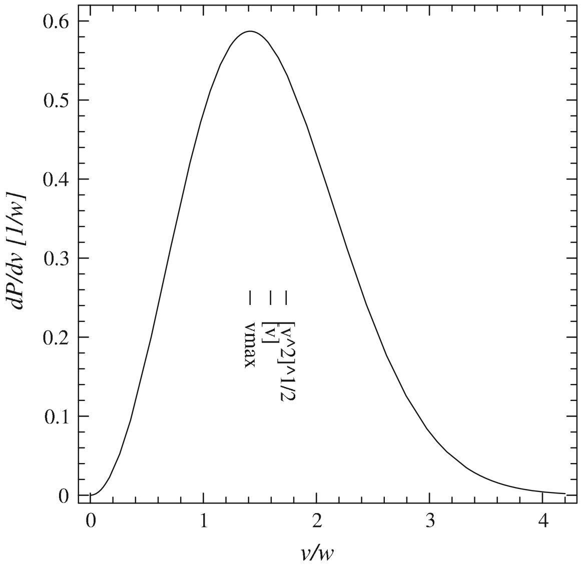


Fig. 4.5 The particle-speed distribution of the high-temperature ideal gas, Eq. (4.51), rescaled by $w = (\beta M)^{-1/2} = (k_\mathrm{B}T/M)^{1/2}$. This distribution peaks at $v_\mathrm{max} = \sqrt{2}\,w \simeq 1.414\,w$, the mean velocity is $[v] = \sqrt{8/\pi}\,w \simeq 1.596\,w$, and the root mean square velocity is $[v^2]^{1/2} = \sqrt{3}\,w \simeq 1.732\,w$

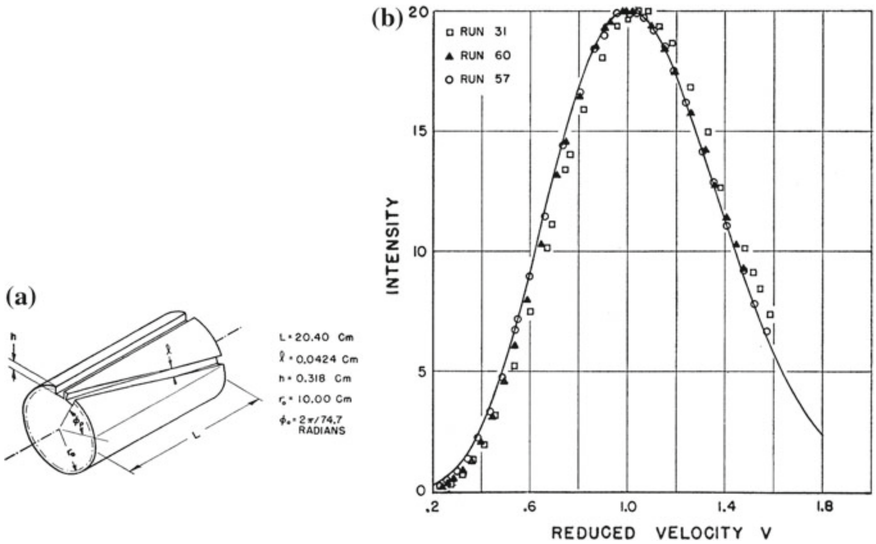


Fig. 4.6 The velocity distribution of the atoms emerging from a oven through a tiny hole was probed by letting them through **a** the spiraling slot of a rotating cylinder, to eventually be counted by a tungsten surface ionization detector. **b** A typical observed distribution of the translational velocities for high-temperature K vapor, compared to the curve of Eq. (4.51), but scaled to $v_{\max}$ rather than w. Reprinted figures with permission from R.C. Miller and P. Kusch, Phys. Rev. **99**, 1314 (1955). Copyright (1955) by the American Physical Society

distribution $P(r)$ of one-electron states in Sect. 2.1.3, should not hide the fact that the most likely velocity is $\mathbf{v} = \mathbf{0}$. Observations of the speed distribution confirm the statistical analysis, see Fig. 4.6.

Internal degrees of freedom affect neither the equation of state nor any of the translational distributions. However, they do contribute an additive temperature-dependent term to the free energy (4.46) [the part $F_{\text{int}}(T) = NF_{1\,\text{int}}(T) = -Nk_{\text{B}}T \ln Z_{1\,\text{int}}$], to the internal energy

$$U = \frac{3}{2}Nk_{\text{B}}T + N\left[F_{1\,\text{int}}(T) - TF'_{1\,\text{int}}(T)\right], \tag{4.52}$$

and therefore also to the heat capacity

$$C_V = \frac{3}{2}Nk_{\text{B}} - NTF''_{1\,\text{int}}(T). \tag{4.53}$$

The internal term vanishes for structureless particles (e.g. electrons), while it contributes significantly to thermodynamics whenever the internal degrees of freedom are associated to excitation energies not too different from $k_{\text{B}}T$. This internal contribution is occasionally relevant for atomic gases and crucial for all molecular gases, as illustrated by the following examples.

4.3.1.2 Internal Degrees of Freedom of Molecules

We report here briefly the example of diatomic molecules, in the approximation that rotational and vibrational motions are independent (Sect. 3.3). The partition function factorizes $Z_{1\,\mathrm{int}} = Z_{1\,\mathrm{vib}}\, Z_{1\,\mathrm{rot}}$, and therefore the free energy $F_{1\,\mathrm{int}} = -k_\mathrm{B} T\ \ln Z_{1\,\mathrm{int}} = F_{1\,\mathrm{rot}} + F_{1\,\mathrm{vib}}$.

For the *rotational* statistics

$$Z_{1\,\mathrm{rot}} = \sum_{l=0}^{\infty} (2l+1) \exp\left(-\frac{\Theta_\mathrm{rot}}{T}\, l[l+1]\right) \tag{4.54}$$

is a function of the dimensionless ratio $\beta\, \hbar^2/(2I) = \Theta_\mathrm{rot}/T$, where the characteristic temperature $\Theta_\mathrm{rot} = \hbar^2/(2I k_\mathrm{B})$. The series of (4.54) cannot be evaluated in a closed form.[4] However, the characteristic temperature Θ_rot is often small (e.g. 85 K for H_2, 15 K for HCl, 2.9 K for N_2). When $\Theta_\mathrm{rot}/T \ll 1$ the exponential in $Z_{1\,\mathrm{rot}}$ varies slowly with l and many terms contribute to the sum in Eq. (4.54): it is a good approximation to replace it with an elementary integral:

$$\begin{aligned} Z_{1\,\mathrm{rot}} &\simeq \int_0^\infty (2l+1) \exp\left[-\frac{\Theta_\mathrm{rot}}{T}\, l(l+1)\right] dl = \int_0^\infty \exp\left(-\frac{\Theta_\mathrm{rot}}{T}\, y\right) dy \\ &= \frac{T}{\Theta_\mathrm{rot}} \qquad [T \gg \Theta_\mathrm{rot}]. \end{aligned} \tag{4.55}$$

The high-temperature rotational contribution to the thermodynamic functions is therefore:

$$F_{1\,\mathrm{rot}} = \frac{F_\mathrm{rot}}{N} \simeq -k_\mathrm{B} T \ln \frac{T}{\Theta_\mathrm{rot}} \tag{4.56}$$

$$U_{1\,\mathrm{rot}} = \frac{U_\mathrm{rot}}{N} \simeq k_\mathrm{B} T \tag{4.57}$$

$$C_{V\,1\,\mathrm{rot}} = \frac{C_{V\,\mathrm{rot}}}{N} \simeq k_\mathrm{B} \tag{4.58}$$

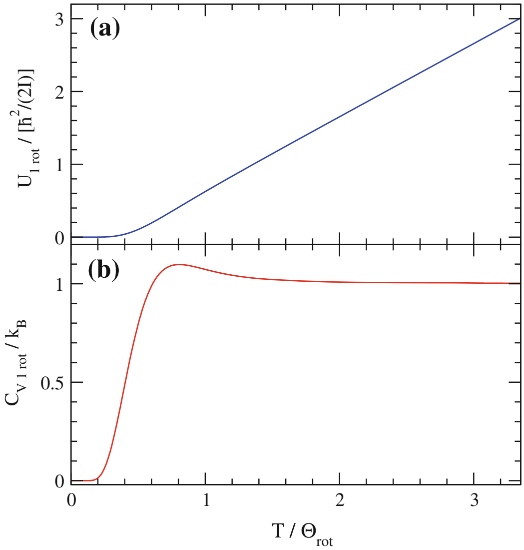
$$S_{1\,\mathrm{rot}} = \frac{S_\mathrm{rot}}{N} \simeq k_\mathrm{B} + k_\mathrm{B} \ln \frac{T}{\Theta_\mathrm{rot}} \tag{4.59}$$

[see Eqs. (4.22)–(4.25)]. At lower temperature $T \sim \Theta_\mathrm{rot}$ instead, truncating the series (4.54) to a finite number of terms $l_\mathrm{max} \gtrsim 2\sqrt{T/\Theta_\mathrm{rot}}$ approximates $Z_{1\,\mathrm{rot}}$ better. Figure 4.7 reports the temperature dependence of the rotational heat capacity and internal energy per molecule. Characteristically, as temperature is raised,

---

[4] Equation (4.54) and its consequences are inaccurate for homoatomic molecules [31], where nuclear spin and indistinguishableness play a role, due to ergodicity violations (e.g. slow ortho–para-hydrogen inter-conversion). In such cases, one must consider even and odd l separately. In the high-T limit, Eqs. (4.56)–(4.59) hold, though.

Fig. 4.7 The temperature dependence of the rotational contribution to **a** the molecular internal energy $U_{1\,\mathrm{rot}}$ and **b** the molecular heat capacity $C_{V\,1\,\mathrm{rot}}$. The energy scale of the problem is set by $\hbar^2/(2I)$, which translates into a temperature scale $\Theta_{\mathrm{rot}} = \hbar^2/(2Ik_B)$. The curves approach the high-temperature limits of Eqs. (4.57) and (4.58), for $T \gg \Theta_{\mathrm{rot}}$



the rotational degree of freedom "unfreezes", reaching the *classical equipartition* limit at large temperature $T \gg \Theta_{\mathrm{rot}}$. Relative equilibrium populations $P_l = (2l+1)\exp(-l[l+1]\Theta_{\mathrm{rot}}/T)\,/Z_{1\,\mathrm{rot}}$ account for the observed relative intensities of the rotational structures in molecular spectra (Figs. 3.9 and 3.10).[5]

As for the *vibrational* statistics in the harmonic approximation, Eq. (3.21), the series is evaluated in closed form, thus yielding an *exact* partition function valid at all temperatures:

$$Z_{1\,\mathrm{vib}} = \sum_{v=0}^{\infty} \exp\left(-\beta\hbar\omega\left[v+\frac{1}{2}\right]\right) = \exp\left(-\frac{x}{2}\right)\sum_{v=0}^{\infty}\exp(-v\,x)$$
$$= \exp\left(-\frac{x}{2}\right)\frac{1}{1-\exp(-x)} = \frac{1}{2\sinh(x/2)}, \quad \text{with } x = \beta\hbar\omega = \frac{\Theta_{\mathrm{vib}}}{T}. \quad (4.60)$$

---

[5] The relative intensities of the rotational lines are proportional to the populations P_l provided that the dipole matrix elements averaged over the initial $(2l+1)$ degenerate states and summed over the final states $l' = l+1$ (R branch) or $l' = l-1$ (P branch) are independent of l. The radial part of the matrix element is, but calculation shows that the angular part equals $(l+1)/(2l+1)$ for the R branch and $l/(2l+1)$ for the P branch. For not too small l, these factors are both close to 0.5, and are therefore often approximated to constants.

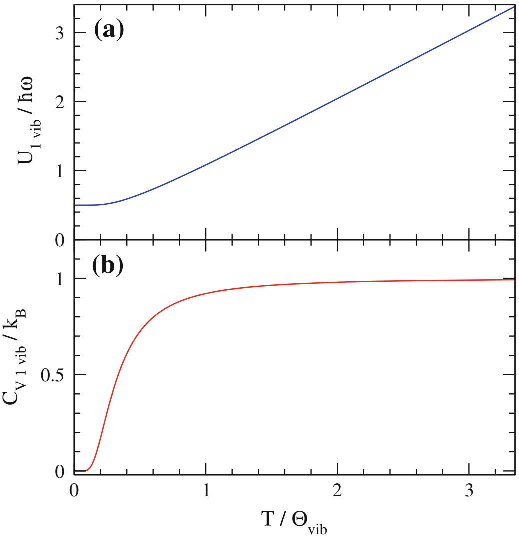


Fig. 4.8 The temperature dependence of **a** the vibrational contribution to the molecular internal energy $U_{1\,\mathrm{vib}}$, Eq. (4.62), and **b** the molecular heat capacity $C_{V\,1\,\mathrm{vib}}$, Eq. (4.63). The energy scale of the problem is set by $\hbar\omega$, which translates into a temperature scale $\Theta_{\mathrm{vib}} = \hbar\omega/k_{\mathrm{B}}$

$Z_{1\,\mathrm{vib}}$ is a function of the dimensionless ratio $x = \Theta_{\mathrm{vib}}/T$, where the characteristic temperature $\Theta_{\mathrm{vib}} = \hbar\omega/k_{\mathrm{B}}$. For the vibrational degree of freedom, the thermodynamic functions of Eqs. (4.22)–(4.25) are:

$$F_{1\,\mathrm{vib}} = \frac{F_{\mathrm{vib}}}{N} = \frac{\hbar\omega}{2} + k_{\mathrm{B}}T\ln\left(1 - e^{-x}\right) \tag{4.61}$$

$$U_{1\,\mathrm{vib}} = \frac{U_{\mathrm{vib}}}{N} = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^x - 1} \tag{4.62}$$

$$C_{V\,1\,\mathrm{vib}} = \frac{C_{V\,\mathrm{vib}}}{N} = k_{\mathrm{B}}\frac{x^2\,e^x}{(e^x - 1)^2} = k_{\mathrm{B}}\left[\frac{x/2}{\sinh(x/2)}\right]^2 \tag{4.63}$$

$$S_{1\,\mathrm{vib}} = \frac{S_{\mathrm{vib}}}{N} = -k_{\mathrm{B}}\ln\left(1 - e^{-x}\right) + k_{\mathrm{B}}\frac{x}{e^x - 1}. \tag{4.64}$$

Figure 4.8 depicts the temperature dependence of the vibrational heat capacity and internal energy per molecule. Similar to rotations, as temperature is raised, the vibrational degree of freedom "unfreezes", reaching the classical limit at $T \gg \Theta_{\mathrm{vib}}$. Note that the high-$T$ limit for *one* oscillator provides a contribution $k_{\mathrm{B}}T$ to U_1 (and thus k_{B} to $C_{V\,1}$) equal to that of the *two* rotational degrees of freedom. The reason is that one harmonic oscillator is associated to two quadratic terms (kinetic and potential) in the Hamiltonian, as many as the two kinetic quadratic terms for the two rotational degrees of freedom.

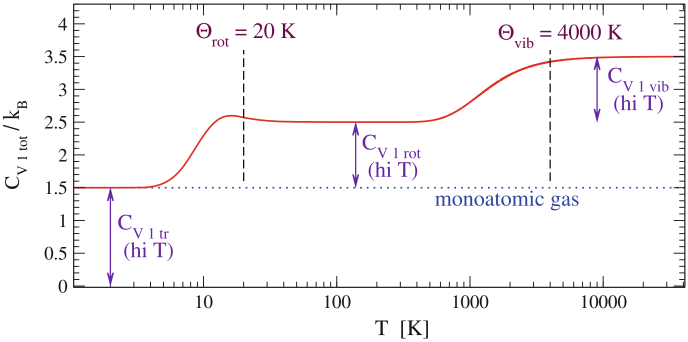


Fig. 4.9 The broad-range temperature dependence of the total heat capacity of a hypothetical diatomic molecule (inspired to HCl), compared to that of a monoatomic gas (*dotted*)

Typical vibrational temperatures of a few diatomic molecules are: 6,300 K for H_2, 4,300 K for $H\,^{35}Cl$, 3,400 K for N_2, 403 K for $K\,^{35}Cl$. With such values of $\Theta_{\rm vib} \gg \Theta_{\rm rot}$, contrary to the rotational unfreezing, the vibrational transition from the quantum-frozen to the classical regime is usually accessible to heat-capacity measurements, which find good accord with Eq. (4.63). Deviations at high temperature are mainly due to the failure of the harmonic approximation.

The translational rotational and vibrational contributions to the molecular heat capacity combine additively. In a real gas three physical mechanisms cause the main deviations from the idealized sketch of Fig. 4.9: (i) inter-molecular interactions would lead to liquid and solid phases at a temperature usually $\Theta_{\rm rot} \lesssim T < \Theta_{\rm vib}$; (ii) even in the absence of inter-molecular interactions, quantum statistical effects (discussed below) would make the low-temperature translational heat capacity deviate from $\frac{3}{2}k_B$; (iii) at high temperature $T \gtrsim 10^4$ K anharmonic and molecular dissociation effects start to affect the heat capacity.

4.3.1.3 Finite Degrees of Freedom (e.g. Spin)

The Boltzmann formalism applies also to the statistics of any set of equal degrees of freedom characterized by a (small) finite number of quantum states. For example, one may wish to evaluate the contributions to thermodynamics of the low-lying electronic and/or magnetic excitations of atomic or molecular gases, accounted for by the internal partition function $Z_{1\,\rm int}$ introduced in Eq. (4.37). One also applies this method to magnetic "impurities" in solids or liquids, as long as they are so dilute that their interactions can be neglected. For any such system involving a finite number of states, the single-atom/molecule/impurity partition function and its derivatives are finite sums.

Here we examine the simple but important case of equal non-interacting magnetic moments in an external field. For brevity, we refer to "spins": in reality the magnetic moment is generally proportional to some *total* angular momentum. An angular momentum J coupled to a total magnetic field $\mathbf{B} = B\hat{\mathbf{z}}$ as in Eq. (2.24) spans a $(2J+1)$-dimensional space of states, a basis of which is labeled by the J_z projection quantum number M_J. The "*spin*" partition function is the sum of $(2J+1)$ terms corresponding to the levels of Eq. (2.59):

$$Z_{1\,\mathrm{spin}} = \sum_{M_J=-J}^{J} \exp\left(-\tilde{\beta}\, M_J\right), \tag{4.65}$$

where the dimensionless ratio $\tilde{\beta} = \beta\, g_j \mu_\mathrm{B} B = \Theta_B/T$ is the inverse temperature scaled by the characteristic $\Theta_B = g_j \mu_\mathrm{B} B/k_\mathrm{B}$ (g_j is the relevant g-factor). It is a simple exercise to evaluate $Z_{1\,\mathrm{spin}}$ for any given spin J.

For the reader's convenience, here we report the results for the simplest case, *spin* $J = 1/2$:

$$Z_{1\,\mathrm{spin}} = 2\cosh\left(\frac{\Theta_B}{2T}\right) = 2\cosh\frac{\tilde{\beta}}{2} \qquad \left[\mathrm{spin}{-}1/2\right]. \tag{4.66}$$

By evaluating Eqs. (4.22)–(4.25) for $Z_{1\,\mathrm{spin}}$, the spin-1/2 thermodynamic functions are:

$$F_{1\,\mathrm{spin}} = \frac{F_\mathrm{spin}}{N} = -k_\mathrm{B} T \ln\left(2\cosh\frac{\tilde{\beta}}{2}\right) \tag{4.67}$$

$$U_{1\,\mathrm{spin}} = \frac{U_\mathrm{spin}}{N} = -\frac{g\mu_\mathrm{B} B}{2}\tanh\frac{\tilde{\beta}}{2} \tag{4.68}$$

$$C_{V\,1\,\mathrm{spin}} = \frac{C_{V\,\mathrm{spin}}}{N} = \frac{k_\mathrm{B}\,\tilde{\beta}^2}{2\cosh(\tilde{\beta}) + 2} \tag{4.69}$$

$$S_{1\,\mathrm{spin}} = \frac{S_\mathrm{spin}}{N} = k_\mathrm{B}\left[\ln\left(2\cosh\frac{\tilde{\beta}}{2}\right) - \frac{\tilde{\beta}}{2}\tanh\frac{\tilde{\beta}}{2}\right]. \tag{4.70}$$

The temperature dependence of the internal energy and heat capacity per spin is drawn in Fig. 4.10. As temperature is raised from absolute zero, like molecular rotations and vibrations, the spin degree of freedom "unfreezes", with an increase in heat capacity. However, when T is raised further, due to the finite spectrum, the internal energy cannot increase indefinitely: $U_{1\,\mathrm{spin}}$ flattens out, thus the spin heat capacity decays to zero at high temperature $T \gg \Theta_B$.

The density of magnetization of a uniform system of magnetic moments

$$\mathbf{M} = \frac{N}{V}\,[\boldsymbol{\mu}_1] \tag{4.71}$$

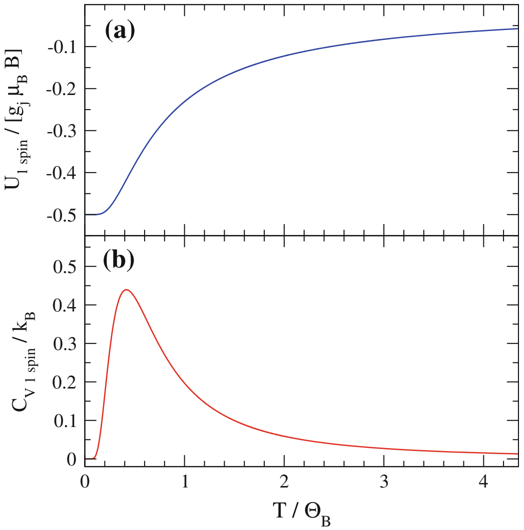


Fig. 4.10 The temperature dependence of **a** the internal energy per spin $U_{1\,\mathrm{spin}}$, Eq. (4.68), and **b** the heat capacity per spin $C_{V\,1\,\mathrm{spin}}$, Eq. (4.69). These quantities characterize the thermodynamics of a magnetic moment in a uniform field B, which sets the energy scale of the problem to $g_j\mu_\mathrm{B}B$, and the corresponding temperature scale $\Theta_B = g_j\mu_\mathrm{B}B/k_\mathrm{B}$. Note that, due to the spectrum being bound from *above*, at high temperature $U_{1\,\mathrm{spin}}$ does not grow indefinitely, but rather approaches a constant (zero)

is a quantity of straightforward experimental accessibility and relevance. In such an ideal system $\mathbf{M}$ is necessarily oriented parallel to the magnetic field, $\mathbf{M} = M\hat{\mathbf{z}}$. We have

$$[\mu_{z\,1}] = \mathrm{Tr}(\hat{\rho}_1 \mu_{z\,1}) = \sum_{M_J=-J}^{J} -g\mu_\mathrm{B} M_J \frac{\exp(-\tilde{\beta} M_J)}{Z_{1\,\mathrm{spin}}}$$
$$= -\frac{1}{B} \sum_{M_J=-J}^{J} \mathscr{E}_{M_J} P_{M_J} = -\frac{U_{1\,\mathrm{spin}}}{B} \tag{4.72}$$

$$M_z = \frac{N[\boldsymbol{\mu}_{z\,1}]}{V} = -\frac{N U_{1\,\mathrm{spin}}}{VB} = \frac{g\mu_\mathrm{B} N}{2V} \tanh \frac{\tilde{\beta}}{2} \qquad [\text{spin-}1/2]\,. \tag{4.73}$$

The average magnetization changes with temperature following the same functional dependence as the internal energy, apart for a constant factor $-N/(VB)$, thus Fig. 4.10a can also be read as magnetization as a function of temperature for the $J = 1/2$ system. For weak field, $\tilde{\beta} \to 0$, the hyperbolic tangent can be expanded to lowest order, obtaining the linear response of the localized spins to a weak total field $\mathbf{B}$:

$$M_z \simeq \frac{g\mu_\mathrm{B}N}{4V}\,\tilde{\beta} = \chi_B\,B, \quad \text{with} \quad \chi_B = \frac{N}{V}\left(\frac{g\mu_\mathrm{B}}{2}\right)^2 \frac{1}{k_\mathrm{B}T} \qquad \left[\text{spin-}1/2\right]. \tag{4.74}$$

χ_B represents the (weak-field) paramagnetic susceptibility. The characteristic inverse-T dependence of the *Curie susceptibility* of free spins reflects the disordering effect of temperature.

In practice it is more common to measure the susceptibility χ_m relative to the *external* applied field strength $H = \varepsilon_0 c^2\,B_\mathrm{ext}$ [in A/m]. The relation between χ_B and χ_m derives from

$$\mathbf{M} = \chi_B\mathbf{B} = \chi_B\,[\mathbf{B}_\mathrm{ext} + \mathbf{B}_\mathrm{int}] = \chi_B\left[(\varepsilon_0 c^2)^{-1}\mathbf{H} + (\varepsilon_0 c^2)^{-1}\mathbf{M}\right], \tag{4.75}$$

where we use the relation $\mathbf{B}_\mathrm{int} = (\varepsilon_0 c^2)^{-1}\,\mathbf{M}$ for the magnetic field of a uniformly magnetized material. We solve Eq. (4.75) for $\mathbf{M}$, obtaining

$$\mathbf{M} = \frac{(\varepsilon_0 c^2)^{-1}\,\chi_B}{1 - (\varepsilon_0 c^2)^{-1}\,\chi_B}\,\mathbf{H}. \tag{4.76}$$

Therefore, the desired expression for the dimensionless susceptibility χ_m is

$$\chi_m = \frac{(\varepsilon_0 c^2)^{-1}\,\chi_B}{1 - (\varepsilon_0 c^2)^{-1}\,\chi_B}. \tag{4.77}$$

The reader should not worry of a diverging denominator in Eq. (4.77): even at a low temperature of 1 K and for $g = 2$, the divergence condition $\chi_B = \varepsilon_0 c^2$ would require a density $N/V \approx 1.3 \times 10^{29}\,\mathrm{m}^{-3}$, far too large for dilute spin-carrying impurities. Actual macroscopic magnetic states are not so much due to a self-sustained magnetization due to a divergence in Eq. (4.77), as to far larger exchange interactions similar to those originating Hund's rules in open-shell atoms, see Sect. 2.2.8.2.

4.3.2 Low-Temperature Fermi and Bose Gases

In Sect. 4.3.1 we discovered that, at high temperature, non-interacting identical fermions or bosons both behave as an ideal classical gas, with the occasional addition of internal degrees of freedom. At low temperature however, spectacular fermion/boson differences show up, as a consequence of the radically different constraints over the occupation of the single-particle states.

The calculation of the exact partition function Z (4.34) at arbitrary T can be carried out by replacing the focus on particles (labeled by i) with a focus in single-particle states (labeled by α). The N sums over the single-particle quantum numbers α_i become sums over the occupation numbers n_α of the single-particle states. In terms of n_α, the total N-particle energy in the exponential $\sum_i^N \mathscr{E}_{\alpha_i}$ is rewritten as $\sum_\alpha n_\alpha \mathscr{E}_\alpha$. In this form, the exponential factorizes as $\exp\left(-\beta \sum_\alpha n_\alpha \mathscr{E}_\alpha\right) = \prod_\alpha \exp(-\beta n_\alpha \mathscr{E}_\alpha)$. The partition function

$$\begin{aligned} Z = \sum_{\alpha_1\, \alpha_2 \ldots \alpha_N} \frac{n_0!\, n_1!\, n_2! \ldots}{N!}\, e^{-\beta \sum_\alpha n_\alpha \mathscr{E}_\alpha} &= \sum_{\substack{\{n_\alpha\} \\ \sum_\alpha n_\alpha = N}} e^{-\beta \sum_\alpha n_\alpha \mathscr{E}_\alpha} \\ &= \sum_{\substack{\{n_\alpha\} \\ \sum_\alpha n_\alpha = N}} \prod_\alpha \left(e^{-\beta \mathscr{E}_\alpha}\right)^{n_\alpha} . \end{aligned} \tag{4.78}$$

The occupancies n_α are 0 or 1 for fermions, and 0, 1, 2, 3,… for bosons. The binomial coefficients correcting the overcounting of the N-particle states by the $\sum_{\alpha_1 \ldots \alpha_N}$ sum of Eq. (4.34), is suppressed in the n_α-sum, because now each N-particle state (identified by its occupation numbers) is counted once only, as it should. Unfortunately, the constraint of fixed total number of particles N makes the sum over the occupancies in Eq. (4.78) extremely difficult to compute: without this constraint, one could exchange sum and product, and the sums would all look alike.

To get rid of the fixed-N constraint we use a trick: we replace the canonical ensemble, where the number of particles is fixed, with the *grand canonical ensemble*, where N is allowed to vary, to describe a thermodynamic system which is (weakly) exchanging not only energy but also particles with the rest of the universe. Figure 4.11 illustrates the need to subtract μN to the energy eigenvalue in the expression (4.11) of the partition function. For an arbitrary system of equal particles, summing the canonical Z over N yields the *grand partition function*

$$Q = \sum_{N=0}^{\infty} \sum_{m(N)} e^{-\beta(E_{m(N)} - \mu N)} = \tilde{\mathrm{Tr}}\left(e^{-\beta(H - \mu \hat{N})}\right), \tag{4.79}$$

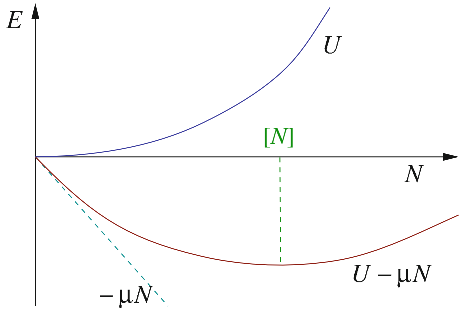


Fig. 4.11 At fixed temperature, the internal energy U is generally a convex function of the number of particles N, which is minimized at some N, e.g. $N = 0$. The addition of $-\mu N$ to U shifts the equilibrium point to some tunable average particle number $[N]$, which is an increasing function of the parameter μ

where $\hat{N}$ is the operator counting the number of particles, and $\tilde{\mathrm{Tr}}$ indicates summing over all states m and all number of particles N. Q plays the same role for the grand canonical ensemble as Z for the canonical ensemble. Detailed analysis, similar to that of Sect. 4.2.1, yields the relation between the grand partition function and thermodynamics. In particular, β identifies with the inverse temperature and μ with the *chemical potential*. We have:

$$J(T, V, \mu) = VP(T, \mu) = k_{\mathrm{B}}T \ln Q(T, V, \mu) \tag{4.80}$$

$$F = \mu\,[N] - J \tag{4.81}$$

$$U = \mu\,[N] - J + TS \tag{4.82}$$

$$G = J + F = \mu\,[N] \tag{4.83}$$

$$[N] = \left.\frac{\partial J}{\partial \mu}\right|_{T,V} = V\left.\frac{\partial P}{\partial \mu}\right|_{T} \tag{4.84}$$

$$\mu = \left.\frac{\partial F}{\partial [N]}\right|_{T,V} = \left.\frac{\partial G}{\partial [N]}\right|_{T,P} \tag{4.85}$$

$$S = \left.\frac{\partial J}{\partial T}\right|_{\mu,V} = V\left.\frac{\partial P}{\partial T}\right|_{\mu}. \tag{4.86}$$

Equation (4.80), the analogous of Eq. (4.22), links Q to the thermodynamic potentials. The listed relations allow us to compute all thermodynamic quantities for a system at equilibrium with a reservoir of energy *and* of particles. See Ref. [30] for further detail.

Armed with this new tool, we proceed to compute the grand partition function Q for a system of noninteracting identical bosons or fermions. Recalling that $N = \sum_\alpha n_\alpha$, we rearrange the exponent of Eq. (4.78) and get rid of the constraint thanks to the new summation over N:

$$\begin{aligned} Q =& \sum_{N=0}^{\infty} \sum_{\substack{\{n_\alpha\} \\ \sum_\alpha n_\alpha = N}} e^{-\beta[(\sum_\alpha n_\alpha \mathscr{E}_\alpha) - \mu N]} = \sum_{N=0}^{\infty} \sum_{\substack{\{n_\alpha\} \\ \sum_\alpha n_\alpha = N}} e^{-\beta[(\sum_\alpha n_\alpha \mathscr{E}_\alpha) - \mu \sum_\alpha n_\alpha]} \\ =& \sum_{\{n_\alpha\}} e^{\sum_\alpha \beta(\mu - \mathscr{E}_\alpha) n_\alpha} = \sum_{\{n_\alpha\}} \prod_\alpha e^{\beta(\mu - \mathscr{E}_\alpha) n_\alpha} \\ =& \prod_\alpha \sum_{n_\alpha} e^{\beta(\mu - \mathscr{E}_\alpha) n_\alpha} = \prod_\alpha \sum_{n_\alpha} \left(e^{\beta(\mu - \mathscr{E}_\alpha)}\right)^{n_\alpha}. \end{aligned}$$

Next, observe that the occupation numbers n_α are mute summation indexes. There is no real reason to tell them apart: just call them all n. Accordingly, we write the grand partition function as

$$Q = \prod_\alpha \left[\sum_n \left(e^{\beta(\mu - \mathscr{E}_\alpha)}\right)^n\right]. \tag{4.87}$$

The sum in square brackets can be carried out explicitly: for fermions ($n = 0, 1$) it equals $1 + e^{\beta(\mu - \mathscr{E}_\alpha)}$; for bosons ($n = 0, 1, 2, 3, \dots$) this sum is a geometric series, whose summation[6] gives $(1 - e^{\beta(\mu - \mathscr{E}_\alpha)})^{-1}$. By introducing a quantity $\theta = +1$ for bosons and $\theta = -1$ for fermions, we can write Q in a form factorized over the single-particle states and valid for both bosons and fermions:

$$Q = \prod_\alpha \left(1 - \theta e^{\beta(\mu - \mathscr{E}_\alpha)}\right)^{-\theta} . \tag{4.88}$$

The grand potential (4.80) provides the connection to thermodynamics:

$$J = PV = k_\mathrm{B} T \ln Q = -\theta k_\mathrm{B} T \sum_\alpha \ln \left(1 - \theta e^{\beta(\mu - \mathscr{E}_\alpha)}\right) . \tag{4.89}$$

This expression, with Eqs. (4.81)–(4.86), determines all the thermodynamics of noninteracting bosons/fermions. For noninteracting free particles, the single-particle states $|\alpha\rangle$ are labeled, as discussed in Sect. 4.3.1, by the single-particle momentum $\mathbf{p}$, plus possibly internal degrees of freedom. At low temperature any nontrivial internal dynamics is usually "frozen". There remain accessible only g_s *degenerate* states labeled by a spin variable m_s (representing for example the $\hat{\mathbf{z}}$ projection of the total angular momentum of the particle). In practice, like in Eq. (4.40), the α-sum represents a sum over n_x, n_y, n_z, and (possibly) m_s. As the translational levels are very dense, the n-sum can be replaced by an integration over energy $\mathscr{E}$, weighted [in the spirit of Eq. (4.12)] by the density of translational states $g_\mathrm{tr}(\mathscr{E})$ of Eq. (4.48):

$$\sum_\alpha \to \sum_{m_s} \sum_{n_x n_y n_z} \to g_s \int_0^\infty d\mathscr{E}\, g_\mathrm{tr}(\mathscr{E}) \to \int_0^\infty d\mathscr{E}\, g(\mathscr{E}) . \tag{4.90}$$

Here we introduce the total density of states $g(\mathscr{E}) = g_s\, g_\mathrm{tr}(\mathscr{E})$. We obtain an equation of state

$$\begin{aligned} P = \frac{J}{V} &= -\theta k_\mathrm{B} T \int_0^\infty d\mathscr{E}\, \frac{g(\mathscr{E})}{V} \ln\left(1 - \theta\, e^{\beta(\mu - \mathscr{E})}\right) \\ &= -\theta k_\mathrm{B} T g_s \frac{(2M)^{3/2}}{4\pi^2 \hbar^3} \int_0^\infty d\mathscr{E}\, \sqrt{\mathscr{E}} \ln\left(1 - \theta\, e^{\beta(\mu - \mathscr{E})}\right) . \end{aligned} \tag{4.91}$$

---

[6] This series converges only if the number $e^{\beta(\mu - \mathscr{E}_\alpha)} < 1$, i.e. only if μ is smaller than the smallest single-particle energy $\mathscr{E}_\alpha$, which usually is 0. Therefore μ must be negative for bosons.

This integral can be rewritten in terms of the dimensionless variable $\tilde{\mathscr{E}} = \mathscr{E}\beta$, and integrated by parts to obtain:

$$P = -\theta (k_B T)^{5/2} g_s \frac{(2M)^{3/2}}{4\pi^2\hbar^3} \int_0^\infty d\tilde{\mathscr{E}} \sqrt{\tilde{\mathscr{E}}} \ln\left(1 - \theta e^{\beta\mu - \tilde{\mathscr{E}}}\right) \tag{4.92}$$

$$= \frac{2}{3} \frac{2}{\sqrt{\pi}} k_B T g_s \Lambda^{-3} \int_0^\infty d\tilde{\mathscr{E}} \frac{\tilde{\mathscr{E}}^{3/2}}{e^{\tilde{\mathscr{E}} - \beta\mu} - \theta},$$

where the thermal length Λ was defined in Eq. (4.43). This equation of state expresses the pressure of a gas of ideal particles in terms of its temperature and chemical potential. This result is unpractical, since μ is a quantity of difficult experimental access: it would be preferable to express P in terms of T and the density $[N]/V$, like in the high-temperature limit Eq. (4.47). Unfortunately, we have no simple analytic expression of μ as a function of the density, thus no convenient explicit equation of state either. More interestingly, by computing the internal energy U through Eq. (4.82), it may be verified that the high-temperature ideal-gas relation

$$P = \frac{2}{3}\frac{U}{V} \tag{4.93}$$

applies at any temperature,[7] for both bosons and fermions, even though from the point of view of the equation of state (4.92), at low temperature ideal bosons and fermions are far from an "ideal gas" in the thermodynamic sense.

It is interesting to examine the leading *high-temperature* correction [30] to the ideal-gas behavior Eq. (4.47) due to quantum statistics:

$$P = \frac{[N]}{V} k_B T \left[1 - \theta\, 2^{-5/2}\, \delta + \mathscr{O}(\delta)^2\right], \quad \text{for } \delta = \frac{\Lambda^3\,[N]}{g_s\, V} = \frac{(2\pi)^{3/2}\,\hbar^3}{(M k_B T)^{3/2}} \frac{[N]}{g_s\, V} \ll 1. \tag{4.94}$$

The explicit form of the *"degeneracy" parameter* δ makes it clear that Eq. (4.94) is a high-temperature and low-density expansion. The sign of the leading correction indicates the opposite tendencies of boson statistics to decrease pressure, and of fermion statistics to increment it. This is a high-temperature hint at the "better social character" of bosons compared to fermions. The experimental verification of these corrections in real gases (e.g. of atoms) is very difficult, because inter-particle

---

[7] We focused the present analysis on the low-temperature regime, the so-called "degenerate" gas. In fact, the only assumption about temperature is that all internal degrees should be either frozen or included in g_s. As long as no other internal degree of freedom plays any relevant role, as e.g. in atomic helium, Eqs. (4.89), (4.92), and (4.93) hold for *arbitrary* temperature.

interactions introduce extra corrections to pressure of the same order as or even larger than those associated to indistinguishableness.

At low temperature, the behavior of ideal bosons and fermions becomes radically different. The average occupation number $[n_\alpha]$ of single-particle states is a clear indicator of these differences. This quantity is obtained as

$$[n_\alpha] = \frac{\left(\sum_n n e^{\beta(\mu-\mathscr{E}_\alpha)n}\right) \cdot \prod_{\alpha'\neq\alpha} \sum_n e^{\beta(\mu-\mathscr{E}_{\alpha'})n}}{\prod_{\alpha''} \sum_n e^{\beta(\mu-\mathscr{E}_{\alpha''})n}} = \frac{\sum_n n e^{\beta(\mu-\mathscr{E}_\alpha)n}}{\sum_n e^{\beta(\mu-\mathscr{E}_\alpha)n}}.$$

This fraction is recognized as $\partial/\partial(\beta\mu)\ln[\sum_n e^{\beta(\mu-\mathscr{E}_\alpha)n}]$. After Eq. (4.87) above, we computed the sum inside the logarithm, and found the result $[1-\theta e^{\beta(\mu-\mathscr{E}_\alpha)}]^{-\theta}$. Therefore, we have

$$\begin{aligned}[n_\alpha] &= \frac{\partial}{\partial(\beta\mu)}\ln[1-\theta e^{\beta(\mu-\mathscr{E}_\alpha)}]^{-\theta} = -\theta\,\frac{\partial}{\partial(\beta\mu)}\ln[1-\theta e^{\beta(\mu-\mathscr{E}_\alpha)}] \\ &= -\theta\,\frac{-\theta e^{\beta(\mu-\mathscr{E}_\alpha)}}{1-\theta e^{\beta(\mu-\mathscr{E}_\alpha)}},\end{aligned}$$

which simplifies to

$$[n_\alpha] = \frac{1}{e^{\beta(\mathscr{E}_\alpha-\mu)}-\theta}. \qquad (4.95)$$

In a single formula, Eq. (4.95) collects the celebrated *Bose-Einstein distribution* of boson occupations

$$[n_\alpha]_B = \frac{1}{e^{\beta(\mathscr{E}_\alpha-\mu)}-1} \qquad (4.96)$$

and *Fermi-Dirac distribution* of fermion occupations

$$[n_\alpha]_F = \frac{1}{e^{\beta(\mathscr{E}_\alpha-\mu)}+1}. \qquad (4.97)$$

The average occupation of each single-particle state is a function uniquely of temperature and of the energy of the state itself.[8] However, the presence of all other particles affects each single-particle occupation distribution through the chemical potential μ, which is a function of the total particle density $[N]/V$ and temperature T.

---

[8] And of no other property. For example, in the absence of any applied magnetic field, energy, and therefore occupation, is independent of m_s: at all temperatures the ideal gas is in a spin-unpolarized nonmagnetic state.

4.3.2.1 Fermions

The statistics of a cold gas of ideal fermions is of fundamental interest for the physics of matter, because conduction electrons in many metals can be approximately described as free non-interacting spin-1/2 fermions. At room temperature the thermal length of electrons is $\Lambda \simeq 4\,\text{nm}$, corresponding to a thermal volume $\Lambda^3 \approx 80\,\text{nm}^3$. Such a large Λ^3, multiplied by the electron density $[N]/V \approx 100\,\text{nm}^{-3}$ of a typical metal, Eq. (4.94) yields a huge degeneracy parameter $\delta \approx 4{,}000$. Electrons in metals are therefore well outside the range of validity of the high-temperature expansion (4.94). Many of their properties can rather be understood in terms of the ideal Fermi-gas model in the opposite low-temperature limit.

The $T = 0$ properties of a Fermi gas are those of the ground state of N free noninteracting fermions[9]: QM prescribes that this state should be the permutation-antisymmetric state obtained by filling the N lowest-energy single-particle levels (i.e. the N/g_s shortest-$|\mathbf{k}|$ plane-wave states), up to some maximum single-particle energy ε_{F} called *Fermi energy*, and leaving all the states above empty, as illustrated in Fig. 4.12. Indeed, in the $\beta \to \infty$ limit Eq. (4.97) predicts precisely that the average occupation turns into a step function of energy:

$$\lim_{T\to 0} [n_\alpha]_F = \begin{cases} 1, & \mathscr{E}_\alpha < \mu \\ 1/2, & \mathscr{E}_\alpha = \mu \\ 0, & \mathscr{E}_\alpha > \mu \end{cases}, \tag{4.98}$$

which thus identifies the chemical potential at $T = 0$ with ε_{F}. Then, by requiring that

$$\begin{aligned} N &= \int_0^\infty d\mathscr{E}\, g(\mathscr{E})\, [n_\alpha]_F = \int_0^\mu d\mathscr{E}\, g(\mathscr{E}) = g_s \frac{(2M)^{3/2} V}{4\pi^2\hbar^3} \int_0^\mu d\mathscr{E}\, \mathscr{E}^{1/2} \\ &= \frac{(2M)^{3/2}\, V\, g_s}{4\pi^2\hbar^3}\, \frac{2}{3}\, \mu^{3/2}, \end{aligned} \tag{4.99}$$

we establish the $T = 0$ relation between the particle density and the chemical potential

$$\varepsilon_{\text{F}} = \mu(T{=}0) = \frac{\hbar^2}{2M} \left(\frac{6\pi^2}{g_s} \frac{N}{V} \right)^{2/3}. \tag{4.100}$$

In $\mathbf{p}$-space, each state within a sphere of radius p_{F} [such that $\varepsilon_{\text{F}} = p_{\text{F}}^2/(2M)$] is filled by g_s fermions, while those outside are empty. To this maximum momentum $p_{\text{F}} = \hbar k_{\text{F}} = \hbar \left(6\pi^2/g_s\, N/V\right)^{1/3}$ (the *Fermi momentum*), there corresponds a maximum

---

[9] Hence, for brevity, we adopt the symbol N for the average number of particles $[N]$.

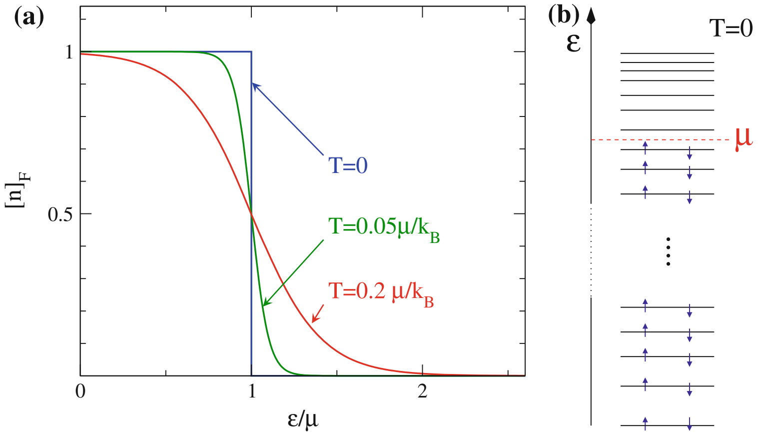


Fig. 4.12 **a** The average filling of the single-particle levels of the ideal Fermi gas as a function of energy $\mathscr{E}$ (measured in units of the chemical potential μ), for three values of temperature. The temperature $0.05\,\mu/k_{\mathrm{B}}$ corresponds to several thousand K for electrons in simple metals. **b** The filling of the single-particle levels of non-interacting fermions (here $g_s = 2$) at $T = 0$

velocity $v_{\mathrm{F}} = p_{\mathrm{F}}/M$. Similarly, the Fermi energy is often expressed in terms of a *Fermi temperature* $T_{\mathrm{F}} = \varepsilon_{\mathrm{F}}/k_{\mathrm{B}}$.

In simple metals, typical densities of conduction electrons ($M = m_e$, $g_s = 2$) of the order $N/V \approx 10^{28} - 10^{29}\,\mathrm{m}^{-3}$ (roughly the inverse cube of typical interatomic separations) yield $\varepsilon_{\mathrm{F}} \approx 2-10\,\mathrm{eV}$, i.e. $T_{\mathrm{F}} \approx 20{,}000-100{,}000\,\mathrm{K}$. This corresponds to $k_{\mathrm{F}} \approx 10^{10}\,\mathrm{m}^{-1}$, $p_{\mathrm{F}} \approx 10^{-24}\,\mathrm{kg\,m\,s}^{-1}$, and a typical *Fermi velocity* $v_{\mathrm{F}} \approx 10^{6}\,\mathrm{m\,s}^{-1}$.

At $T = 0$ it is also straightforward to obtain the internal energy:

$$U = \int_0^\infty d\mathscr{E}\,\mathscr{E}\,g(\mathscr{E})\,[n_\alpha]_F = \int_0^{\varepsilon_{\mathrm{F}}} d\mathscr{E}\,\mathscr{E}\,g(\mathscr{E}) = \frac{(2M)^{3/2} V\, g_s}{4\pi^2\hbar^3} \int_0^{\varepsilon_{\mathrm{F}}} d\mathscr{E}\,\mathscr{E}^{3/2}$$
$$= \frac{(2M)^{3/2} V\, g_s}{4\pi^2\hbar^3}\,\frac{2}{5}\,\varepsilon_{\mathrm{F}}^{5/2}. \qquad (4.101)$$

U can be expressed in terms of the density N/V by substituting Eq. (4.100). If in Eq. (4.101) we substitute just a factor $\varepsilon_{\mathrm{F}}^{3/2}$, we obtain the easier-to-remember relation:

$$U(T=0) = \frac{3}{5} N\,\varepsilon_{\mathrm{F}}. \qquad (4.102)$$

As a special case of Eq. (4.93), the $T = 0$ pressure

$$P(T=0) = \frac{2}{5}\,\frac{N}{V}\,\varepsilon_{\mathrm{F}} \qquad (4.103)$$

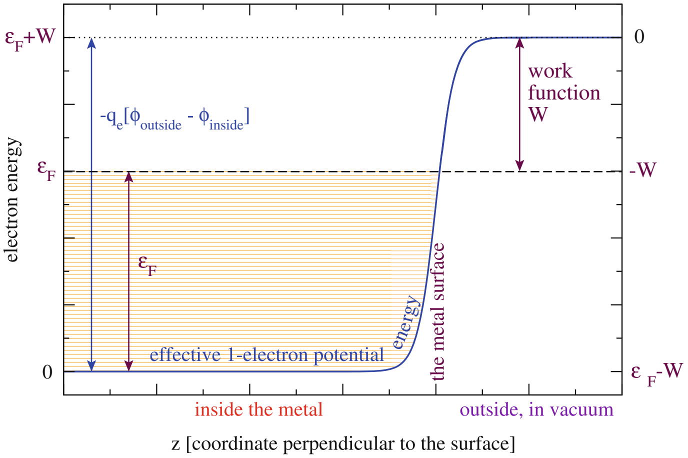


Fig. 4.13 The level scheme of the $T = 0$ Fermi-gas model applied to a metal. The energy scales at the *left* and at the *right* are equally valid. Energy levels below the chemical potential (*dashed line*) are fully occupied states inside the metal. Electrons promoted (e.g. by electromagnetic radiation) to levels above the *dotted line* are free to leave the metal

can be very large, due to the huge average kinetic energy of the fermions, obliged by Pauli's principle to occupy different, orthogonal, plane-wave states. When the $T = 0$ free-electron gas model is applied to a simple metal, one discovers that the pressure exerted by the electron gas is as large as $P \approx 1 - 10\,\text{GPa}$! A sharp potential-energy step at the metal surface, related to the attractive atomic nuclei, maintains this pressure by preventing the conduction electrons from escaping the solid, see Fig. 4.13.

The $T = 0$ properties of the free-electron gas account for several properties of the electrons in simple metals, even at room temperature.[10] The level scheme of Fig. 4.13 illustrates the application of the Fermi-gas model to metals. Note the role played by the work function, namely the minimum energy required to extract an electron from the solid.

We need to extend our analysis to the finite-T regime, in order to determine those thermodynamic quantities that involve temperature explicitly, such as the heat capacity. In the $T \ll T_F$ limit, a systematic expansion of the equation of state (4.92) and the relation (4.85) connecting μ to N/V provides the thermal properties at lowest-order in T. The mathematical details of this procedure [10] (called Sommerfeld expansion) are slightly intricate, but the qualitative trends are straightforward. For example, one can estimate the leading temperature dependence of μ and U by observing that when a small temperature is turned on, the average occupation $[n_\alpha]_F$ changes slightly as

---

[10] The accord of the free-electron model with experimental data of many simple metals is surprisingly good despite neglecting the Coulomb interactions between electrons. The reason is that the electron gas screens the long-range Coulomb forces efficiently. An experimentally observed phenomenon directly related to electron-electron Coulomb repulsion is that of *plasmon* collective excitations, which however occur at rather high energy (few eV), and are therefore of little importance to thermodynamics at ordinary temperature.

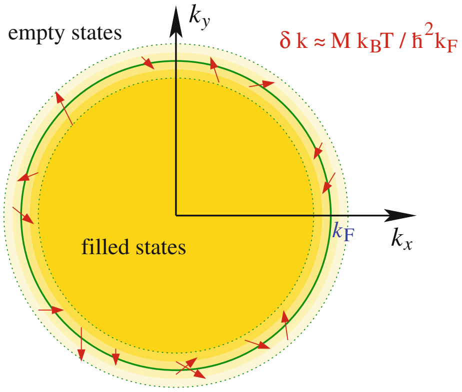


Fig. 4.14 Thermal excitations/deexcitations across the Fermi sphere involve mainly the occupations of the states within a skin of thickness δk immediately inside and outside the Fermi sphere. The skin thickness is greatly exaggerated here compared to the smearing induced by a realistic temperature $T \ll T_F$ of electrons in ordinary metals at room conditions

sketched in Fig. 4.12a from the $T = 0$ step function, providing some weak probability that states above μ are populated at the expense of states below μ. As the density of states $g(\mathscr{E}) \propto \mathscr{E}^{1/2}$ is slightly larger above ε_F than below, to conserve the fermion number $N = \int_0^\infty g(\mathscr{E})\,[n_\alpha]_F\, d\mathscr{E}$ the chemical potential decreases slowly as T increases.[11] The internal energy U increases mainly due to the few electrons moving up from states of a skin region of thickness $\approx k_B T$ below ε_F into states $\approx k_B T$ above (Fig. 4.14): the energy of each excited electron increases by $\approx k_B T$. The number of excited electrons is of the order of the density of states $g(\varepsilon_F)$ times the energy interval $k_B T$ where excitation occurs. The total internal energy increases therefore by approximately $g(\varepsilon_F)(k_B T)^2$. The detailed expansion yields

$$\begin{aligned} U &= U(T=0) + \frac{\pi^2}{6}\, g(\varepsilon_F)\,(k_B T)^2 + \cdots \\ &= \frac{3}{5} N\,\varepsilon_F \left[1 + \frac{5\pi^2}{12}\left(\frac{T}{T_F}\right)^2 + \cdots\right] \quad [T \ll T_F], \end{aligned} \tag{4.104}$$

where the second equality relies on the useful relation valid for spin-1/2 free fermions

$$g(\varepsilon_F) = \frac{3N}{2\varepsilon_F}. \tag{4.105}$$

Derivation of Eq. (4.104) w.r.t. T yields the heat capacity of the ideal Fermi gas:

$$C_V = N\,k_B\,\frac{\pi^2}{2}\,\frac{T}{T_F} + \cdots \qquad [T \ll T_F]. \tag{4.106}$$

---

[11] The low-T dependence of the chemical potential is $\mu = \varepsilon_F\left[1 - \pi^2/12\ (T/T_F)^2 + \cdots\right]$, for $g_s = 2$.

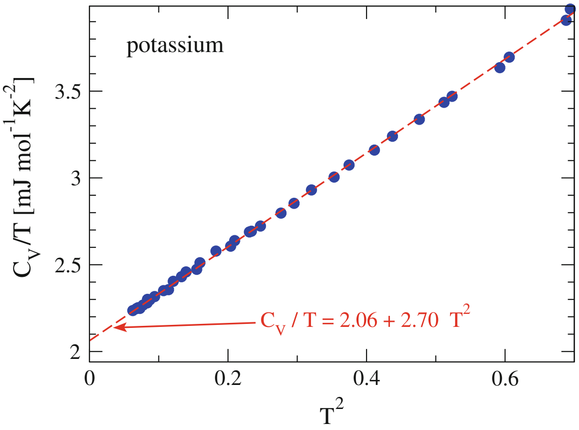


Fig. 4.15 The low-temperature measured molar heat capacity of metallic potassium divided by T, as a function of T^2. For $C_V \simeq a T^1 + b T^3$, the finite intercept at $T^2 = 0$ of the C_V/T curve measures the coefficient a of the T^1 (electronic) contribution; the slope of the graph measures the coefficient b of T^3, which is the contribution of lattice vibrations, see Sect. 5.3.2 (Data from Table 1 of Ref. [32]—the point at $T = 0.4231$ K was suppressed)

Note that compared to the high-temperature ideal-gas value $3/2 N k_B$, the heat capacity is suppressed by a factor $\propto T/T_F$. The reason is that only few fermions with energy very close to the chemical potential can be involved in thermal excitations, the large majority remaining “frozen” in deeper filled states. Experimentally, a T-linear contribution to the total heat capacity is observed in solid metals at low temperature (Fig. 4.15), where lattice-vibration contributions are small, see Sect. 5.3. The electron gas is the responsible for this T-linear contribution: indeed the observed T-linear coefficient for, e.g., potassium $\simeq 2.1\,\text{mJ}\,\text{mol}^{-1}\text{K}^{-2}$ (Fig. 4.15) agrees fairly with the free-electron estimate $N_A\, k_B\, \pi^2/(2T_F) \simeq 1.7\,\text{mJ}\,\text{mol}^{-1}\text{K}^{-2}$ [consistent with the Fermi energy $\varepsilon_F \simeq 2.1$ eV obtained through Eq. (4.100) for the experimental density $N/V \simeq 1.3 \times 10^{28}\,\text{m}^{-3}$ of conduction electrons in potassium].

The Fermi gas is nonmagnetic, as both spin states are equally occupied. Within linear response, the application of an external magnetic field strength $\mathbf{H} = H_z\hat{\mathbf{z}}$ produces a total magnetic field $\mathbf{B} = B_z\hat{\mathbf{z}}$, that we assume to couple only to the spin degrees of freedom: the g_s degeneracy is lifted, and the ideal Fermi gas magnetizes. $\mathbf{M} = M_z\hat{\mathbf{z}}$ denotes the volume density of magnetic moment, like in Eq. (4.71). For spin-1/2 electrons ($g_s = 2$), the magnetization is $M_z = -\mu_B[N_\uparrow - N_\downarrow]/V$. By computing the linear response to the external field, one obtains the magnetic behavior of the ideal Fermi gas. At high T, according to Eq. (4.74), independent spins produce a Curie susceptibility $\chi_B \propto N/T$, while at small $T \ll T_F$ the Pauli principle freezes out the spins of most electrons, the paired ones deep inside the Fermi sphere (Fig. 4.14). Only the approximately $g(\varepsilon_F)\, k_B T$ electrons near the Fermi surface do spin-polarize, thus producing a characteristically T-independent magnetization. Detailed calculation yields:

$$M_z = \frac{\mu_B^2 g(\varepsilon_F)}{V} B_z = \frac{3}{2}\frac{\mu_B^2}{\varepsilon_F}\frac{N}{V} B_z = \frac{3^{1/3}}{4\pi^{4/3}}\frac{q_e^2}{m_e}\left(\frac{N}{V}\right)^{1/3} B_z \qquad [T \ll T_F], \qquad (4.107)$$

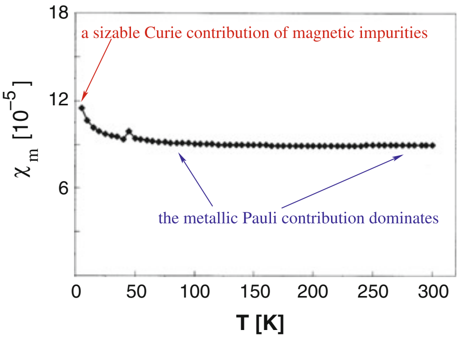


Fig. 4.16 The measured magnetic susceptibility of $Zr_2V_6Sb_9$. Conduction electrons are responsible for the T-independent Pauli susceptibility χ_m of the metal. The raise of χ_m at very small temperature is due to the Curie-type contribution, Eq. (4.74), of magnetic impurities in the sample (Data from Ref. [33])

where the last expression shows the explicit density dependence. Accordingly, the spin susceptibility of the low-temperature ideal spin-1/2 Fermi gas is given by Eq. (4.77), with

$$(\varepsilon_0 c^2)^{-1}\,\chi_B = \frac{3^{1/3}}{4\pi^{4/3}\,\varepsilon_0 c^2}\,\frac{q_e^2}{m_e}\left(\frac{N}{V}\right)^{1/3} = \alpha^2 a_0\left(\frac{3\,N}{\pi\,V}\right)^{1/3}. \tag{4.108}$$

As $(V/N)^{1/3}$ is a few Bohr radii, the magnetic susceptibility $\chi_m \simeq (\varepsilon_0 c^2)^{-1}\,\chi_B$ is tiny, $\approx 10^{-5}$. The reason is that χ_B gathers contributions of the electrons near the Fermi level only, so that it is proportional to $(N/V)^{1/3}$, compared to linear in N/V as for isolated spins, Eq. (4.74). This weak T-independent paramagnetic response (*Pauli paramagnetism*) is a characteristic signature of metals in the experimental study of the magnetic response of materials (see e.g. Fig. 4.16).

4.3.2.2 Bosons

The ground state of N non-interacting bosons is far simpler than that of N fermions: they all occupy the lowest-energy state $\mathbf{k} = \mathbf{0}$ of energy $\mathscr{E} = 0$. If a spin degeneracy g_s is present, the average number of bosons of each spin projection is N/g_s. This situation is reflected in the expression (4.96) for $[n_\alpha]_B$: for $T \to 0$, the occupancies of all positive-energy states vanish, and $\mu \simeq -k_B T/N \to 0^-$ in such a way to ensure that the occupation of the $\mathscr{E} = 0$ state approaches N. Unfortunately, the density of states (4.48) vanishes at $\mathscr{E} = 0$: this indicates that the conversion (4.90) of the sum over the discrete one-particle states into an energy integral is missing completely the most populated state, which becomes dominant at low temperature. A correct treatment, including explicitly the population of the $\mathscr{E} = 0$ state, shows a phase transition at a finite temperature

$$T_{\rm BE} = \frac{6.625}{k_{\rm B}} \frac{\hbar^2}{2M} \left(\frac{N}{V}\right)^{2/3} , \tag{4.109}$$

signaled by the macroscopic filling of the $\mathscr{E} = 0$ level at $T \le T_{\rm BE}$.[12] This low-temperature collective state is called a "*Bose-Einstein condensate*".

Even though many atoms and molecules are bosons, they all turn solid at a much higher temperature than the $T_{\rm BE}$ appropriate for their ordinary densities. The only exception in ordinary matter is ^{4}He, and indeed a *superfluid* transition similar to the Bose-Einstein condensation of an ideal boson gas is observed in ^{4}He near 2 K at ordinary pressure. However, ^{4}He at low temperature is a liquid rather than a gas, indicating that inter-particle interactions play an important role, and the helium fluid can hardly be regarded as an ideal system: more sophisticated tools are necessary to understand the actual nature of the superfluid transition of ^{4}He.

Since the mid 1990s, ultracold droplets of atoms are being produced and kept in a metastable gaseous state inside electromagnetic traps. Progress in cooling techniques, down to the sub-μK range, permits to cross routinely $T_{\rm BE}$ with droplets of boson atoms [34]. In these droplets atoms are much more dilute than in a liquid, thus the atom-atom interactions play a far weaker role. Therefore these droplets provide experimental realizations of Bose-Einstein condensates more similar to the ideal Bose gas than the ^{4}He superfluid state.

In addition to proper material bosons, also the thermodynamic properties of fictitious particles related to harmonic oscillators are described by the Bose-Einstein distribution. Equation (3.21) yields the eigenvalues of an harmonic oscillator in terms of the quantum number v, which counts the nodes of the corresponding wavefunction. Wherever several harmonic oscillators of frequencies ω_α are present, as in a polyatomic molecule, the vibrational state is labeled by all the v_α quantum numbers, and the associated total energy is

$$E_{\rm vib}(v_1, v_2, \ldots) = \sum_\alpha \left(v_\alpha + \frac{1}{2}\right) \hbar\omega_\alpha . \tag{4.110}$$

Compare Eq. (4.110) to the expression

$$E(n_1, n_2, \ldots) = \sum_\alpha n_\alpha \mathscr{E}_\alpha \tag{4.111}$$

used, e.g. in Eq. (4.34), for the energy of noninteracting particles in terms of the occupation numbers n_α of the single-particle states of energy $\mathscr{E}_\alpha$. Apart from an

---

[12] The average occupation of each state $|\alpha\rangle$ follows Eq. (4.95). If one doubles the system size (both N and V), then the average occupation of $|\alpha\rangle$ does not change: it is the density of states $g_{\rm tr}(\mathscr{E}_\alpha)$ that doubles to take care of the extra particles, see Eq. (4.48). Below $T_{\rm BE}$, the $\mathscr{E} = 0$ state marks an exception: its occupation is a finite fraction of N, thus, when doubling the system size, its occupation also doubles.

irrelevant constant zero-point shift $\frac{1}{2}\sum_\alpha \hbar\omega_\alpha$, expressions (4.110) and (4.111) are identical, provided that the following identifications are made:

$$\begin{aligned}
\text{individual oscillator } \alpha &\longleftrightarrow \alpha \text{ single}-\text{particle state}\\
\text{oscillator quantum number } v_\alpha &\longleftrightarrow n_\alpha \text{ occupation number of state } \alpha\\
\text{oscillator energy quantum } \hbar\omega_\alpha &\longleftrightarrow \mathscr{E}_\alpha \text{ single}-\text{particle energy of state } \alpha\\
\text{oscillator eigenvalue } v_\alpha\,\hbar\omega_\alpha &\longleftrightarrow n_\alpha\,\mathscr{E}_\alpha \text{ energy of } n \text{ particles in state } \alpha.
\end{aligned}$$

As $v_\alpha = 0, 1, 2, 3, \ldots$, this identification makes sense for boson occupations. The equality of the spectrum causes a completely equivalent statistical behavior. In detail, consider the average vibrational energy of one oscillator, Eq. (4.62) and remove the zero-point term $\hbar\omega_\alpha/2$. Individual energy levels $v\,\hbar\omega_\alpha$ are proportional to the quantum number v. Thus the average energy $\hbar\omega_\alpha/(\exp x - 1)$, reflects an average value of v, which is $[v] = 1/(\exp x - 1)$. This expression coincides with the average occupation $[n_\alpha]_B$, Eq. (4.96), of a single-boson state $|\alpha\rangle$, provided that $\mathscr{E}_\alpha = \hbar\omega_\alpha$ and $\mu = 0$. One is then lead to think of each step in the harmonic ladder as a fictitious boson-type particle, called "*phonon*" (sound particle) or "*photon*" (light particle), depending on the nature of the involved oscillator. Accordingly, an oscillator in its $|v = 4\rangle$ state is said to carry 4 photons, while an oscillator in its ground state $|v = 0\rangle$ holds no photons. In this way, the Bose-Einstein distribution can be profitably employed to describe the thermodynamics of a set of harmonic oscillators, by replacing the single-particle energy $\mathscr{E}_\alpha$ with the relevant $\varepsilon = \hbar\omega_\alpha$. The lack of chemical potential μ indicates that the average number N of bosons is impossible to control. Contrary to material particles in a vessel, phonons/photons are not conserved: N varies widely as a function of temperature.

It is now straightforward to obtain the statistical properties of a set of independent oscillators. We summarize here the main results for the "photon gas", i.e. the thermodynamics of the normal modes of the electromagnetic fields at thermal equilibrium inside an isothermal cavity. This system is also known as *blackbody radiation*.

The components of the electromagnetic fields in vacuum obey a (Laplace) stationary equation formally identical to the Eq. (B.52) of Schrödinger particles, but with c^2 in place of $\hbar^2/(2m)$ and ω^2 in place of $\mathscr{E}$. The resulting *dispersion relation* $\varepsilon(\mathbf{p})$ [or, equivalently, $\omega(\mathbf{k})$] is

$$\varepsilon(\mathbf{p}) = c|\mathbf{p}| \qquad \text{or} \qquad \omega(\mathbf{k}) = c|\mathbf{k}|, \tag{4.112}$$

(c is the speed of light) rather than

$$\mathscr{E}(\mathbf{p}) = \frac{|\mathbf{p}|^2}{2m} \qquad \text{or} \qquad \omega(\mathbf{k}) = \frac{\hbar|\mathbf{k}|^2}{2m} \tag{4.113}$$

as appropriate for particles of matter with mass m. While the $\omega(\mathbf{k})$ relation is quite different, the equation being structurally the same, solutions are also plane waves. Under the same periodic boundary conditions, the allowed values of wave vector are connected to the box size by the same Eq. (B.55). The "spin" degeneracy of photons is $g_s = 2$, corresponding to the two independent transverse polarizations. By counting the states within a **k**-sphere one obtains the total density of oscillator energies

$$g_s\, g_{\mathrm{ph}}(\varepsilon) = \frac{V}{\pi^2 \hbar^3 c^3}\, \varepsilon^2, \tag{4.114}$$

to be compared with the same counting for Schrödinger particles, Eq. (4.48).

By executing the sum over the independent harmonic oscillators, i.e. integrating over their energy $\varepsilon = \hbar\omega$, one obtains the following thermodynamic relations:

$$\begin{aligned} U &= V \int_0^\infty u(\varepsilon, T)\, d\varepsilon = V \frac{\pi^2 (k_{\mathrm{B}} T)^4}{15\, \hbar^3 c^3}, \\ &\text{with } u(\varepsilon, T) = \frac{1}{V}\, g_s\, g_{\mathrm{ph}}(\varepsilon)\, [n_\varepsilon]_B\, \varepsilon = \frac{1}{\pi^2 \hbar^3 c^3} \frac{\varepsilon^3}{e^{\varepsilon/k_{\mathrm{B}}T} - 1}, \end{aligned} \tag{4.115}$$

$$[N] = \int_0^\infty g_s\, g_{\mathrm{ph}}(\varepsilon) [n_\varepsilon]_B\, d\varepsilon = V \frac{1}{\pi^2 \hbar^3 c^3} (k_{\mathrm{B}} T)^3 \int_0^\infty \frac{y^2}{e^y - 1} dy = V\, \frac{2\, \xi(3)}{\pi^2 \hbar^3 c^3}\, (k_{\mathrm{B}} T)^3, \tag{4.116}$$

where ξ is the Riemann function [$\xi(3) \simeq 1.20206$]. These results were first derived by M.K.E.L. Planck to interpret the experimental data of thermal-radiation spectra. Equation (4.115) makes an important prediction for the spectral density $u(\varepsilon, T)$ (energy per unit volume and energy interval) of equilibrium radiation.[13]

It is more straightforward to probe experimentally the spectral irradiance R (radiated power per unit surface and spectral energy interval) of some radiation than its energy density. As it turns out, R and u are proportional:[14]

$$R(\varepsilon, T) = \frac{c}{4}\, u(\varepsilon, T). \tag{4.117}$$

---

[13] The same quantities are occasionally quoted in terms of frequency ν or wavelength λ, rather than photon energy $\varepsilon = 2\pi\hbar\nu = 2\pi\hbar c/\lambda$. For example: $\tilde{g}(\nu, T) = 8\pi V \nu^2/c^3$, $U = V \int_0^\infty \tilde{u}(\nu, T)\, d\nu$, with $\tilde{u}(\nu, T) = 16\pi^2\, \hbar c^{-3}\, \nu^3 / \{\exp[(2\pi\hbar\nu)/(k_{\mathrm{B}}T)] - 1\}$. And, likewise, $U = V \int_0^\infty \bar{u}(\lambda, T)\, d\lambda$, with $\bar{u}(\lambda, T) = 16\pi^2\, \hbar c\, \lambda^{-5} / \{\exp[(2\pi\hbar c)/(\lambda k_{\mathrm{B}} T)] - 1\}$. Despite expressing the same property, the physical dimension of these spectral densities are of course different: $[u] = \text{Length}^{-3}$, $[\tilde{u}] = \text{Energy} \times \text{Length}^{-3} \times \text{Time}$, $[\bar{u}] = \text{Energy} \times \text{Length}^{-4}$.

[14] The proportionality factor $c/4$ between irradiance R and energy density u originates from the fact that each photon carries its energy at the speed of light c. Moreover, given an infinitesimal surface, only a fraction $(4\pi)^{-1} \int_0^1 \cos\theta\, d\cos\theta\, \int_0^{2\pi} d\varphi = 1/4$ of the photons in the surroundings of that surface crosses it in a given direction.

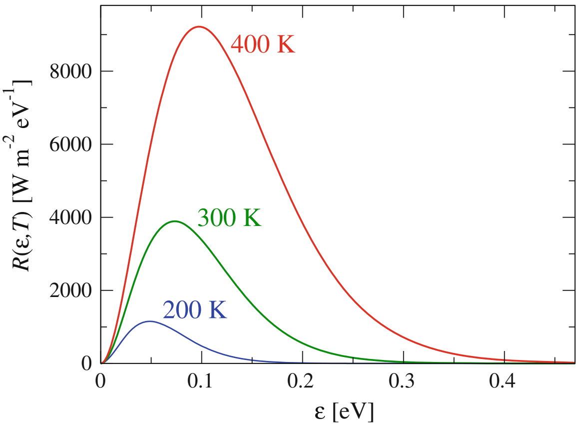


Fig. 4.17 The spectral irradiance $R(\varepsilon, T) = \frac{c}{4}u(\varepsilon, T)$ of the equilibrium (*blackbody*) electromagnetic fields at three temperatures. Note that as T increases, the area under the *curve* increases very rapidly, Eq. (4.118), and its maximum shifts to higher energy (shorter wavelength)

Figure 4.17 reports the spectral irradiance $R(\varepsilon, T)$ as a function of energy. The area under each curve equals the total radiated power per unit surface

$$\int_0^\infty R(\varepsilon, T)\, d\varepsilon = \frac{\pi^2 (k_B T)^4}{60\, \hbar^3 c^2}. \tag{4.118}$$

This result is know as the *Stefan-Boltzmann law*. According to Eq. (4.118), e.g., one square meter of blackbody surface[15] radiates a total 459 W at 300 K. The energy at which a maximum irradiance $R(\varepsilon, T)$ is observed shifts linearly with T (*Wien displacement law*).

The spectral distribution of energy density (and, equivalently, radiative power) of electromagnetic fields at equilibrium is a universal function of temperature (called *blackbody spectrum*), and does not depend on the precise way this equilibrium is established (for example on the optical properties of the material of the cavity enclosing the fields). The distribution (4.115) agrees perfectly with the spectral analysis of radiation escaping from an isothermal cavity through a tiny hole. This agreement is not surprising, because a description of the electromagnetic fields in terms of harmonic oscillators is basically exact. A spectacular realization (Fig. 4.18) of the thermal-equilibrium radiation is that of the cosmic microwave background, a “fossil” relic of an early stage of our universe when it was all at thermal equilibrium.

---

[15] The name “blackbody” to indicate the spectrum of equilibrium radiation originates from the fact that a perfectly absorbing (0 % reflectivity over the entire spectral range of Fig. 1.2) surface irradiates precisely the spectrum $R(\varepsilon, T)$, where T is the temperature of the surface itself. The reason is the detailed energy balance of the fluxes of incoming and outgoing radiation between the surface and the fields, which would be necessarily established if the surface was to be a part of the inside face of an isothermal cavity containing the equilibrium electromagnetic fields.

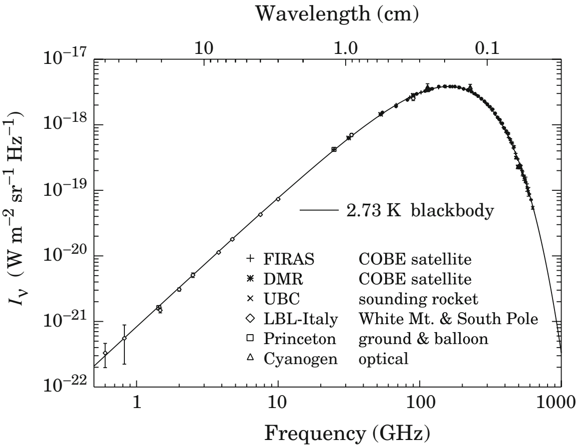


Fig. 4.18 The observed cosmic microwave background frequency spectrum, compared to a 2.73 K blackbody spectrum derived from Eq. (4.115). Note the $\propto \varepsilon^2$ increase at small frequency, and the rapid (exponential) decay past the maximum (From G.F. Smoot and D. Scott, Cosmic Background Radiation, http://www.astro.ubc.ca/people/scott/cbr_aph_00.ps)

In Sect. 5.3.2, we will apply the same statistics, with a modified density of states, to describe the thermal properties of the vibrations of solids (phonons in the Debye model).

4.4 Interaction Matter-Radiation

In Sect. 2.1.9 we have sketched the basic result for the rate—Eq. (2.45)—of spontaneous decay of an atom (or a molecule) from an excited state, as in an emission experiment. In the absence of any external stimulation, the material system decays to a lower-energy state transferring its excess energy to the electromagnetic radiation field. In this section we examine in some detail how an external *stimulating radiation*, as e.g. in an absorption experiment, affects this interaction. First off, this radiation needs to resonate with the energy of a transition between two eigenstates of the system. For simplicity, we shall ignore off-resonance transitions, occurring at much smaller rates.

Consider an ensemble of noninteracting quantum "particles" (say atoms in gas phase) and focus on two single-system levels only: $|1\rangle$ and $|2\rangle$, of energies $\mathcal{E}_1$ and $\mathcal{E}_2 > \mathcal{E}_1$, with instantaneous populations n_1 and n_2, respectively. Resonant radiation of energy $\varepsilon = \hbar\omega = \mathcal{E}_2 - \mathcal{E}_1$ can induce transitions between these two energy levels. In particular, the excitation $|1\rangle \to |2\rangle$ is driven by the presence of radiation. Accordingly, the probability per unit time (i.e. the rate) that an atom initially in state $|1\rangle$ gets excited to the upper level $|2\rangle$ is proportional to the spectral energy density (per unit volume and spectral interval) $\rho(\varepsilon)$ of the electromagnetic field at the resonant energy:

$$R_{1\to 2} = B_{12}\,\rho(\varepsilon), \tag{4.119}$$

where B_{12} is a suitable proportionality constant, depending on the microscopic characteristics of the system and its coupling to the field. Here we neglect nonlinear effects $\mathscr{O}(\rho^2)$. On the other hand, an atom initially in state $|2\rangle$ has a probability per unit time A_{21} to decay spontaneously to state $|1\rangle$ [in the dipole approximation, A_{21} equals the γ_{21} of Eq. (2.45)], plus an additional probability of downward transitions promoted by the presence of stimulating radiation, whose rate is then proportional to the same resonant spectral density:

$$R_{2\to 1} = A_{21} + B_{21}\,\rho(\varepsilon), \tag{4.120}$$

where again B_{21} is a yet-to-determine constant of proportionality.

At any given time, the total number of particles undergoing the $|1\rangle \to |2\rangle$ transition is $n_1 R_{1\to 2}$, and the total number of particles going $|2\rangle \to |1\rangle$ is $n_2 R_{2\to 1}$. Relations (4.119) and (4.120) hold under arbitrary radiation conditions, for example when the ensemble is probed by a radiation beam in an absorption experiment (Fig. 1.3). In particular, these relations hold also when the ensemble and the radiation field are *at equilibrium* at a given temperature. At equilibrium the average populations of individual states remain constant, and this implies that the total number of $|1\rangle \to |2\rangle$ and $|2\rangle \to |1\rangle$ transitions must, on average, be equal:

$$[n_1]\,R_{1\to 2} = [n_2]\,R_{2\to 1} \qquad \text{[at equilibrium]}. \tag{4.121}$$

We substitute Eqs. (4.119) and (4.120) in the balance equation (4.121)

$$[n_1]\,B_{12}\,\rho(\varepsilon) = [n_2]\,\left(A_{21} + B_{21}\,\rho(\varepsilon)\right). \tag{4.122}$$

and solve for $\rho(\varepsilon)$ obtaining

$$\rho(\varepsilon) = \frac{\frac{A_{21}}{B_{21}}}{\frac{[n_1]}{[n_2]}\frac{B_{12}}{B_{21}} - 1}. \tag{4.123}$$

At equilibrium, the ratio of the populations equals the probability ratio which, according to Boltzmann statistics, is simply $[n_1]/[n_2] = P_1/P_2 = \exp(\beta[\mathscr{E}_2 - \mathscr{E}_1]) = \exp(\beta\varepsilon)$. At equilibrium, the spectral density of the radiation field follows the universal energy dependence described in Sect. 4.3.2.2, in particular by Eq. (4.115): $\rho(\varepsilon) = u(\varepsilon, T)$. Accordingly,

$$\frac{A_{21}}{B_{21}}\,\frac{1}{\exp\left(\varepsilon/k_{\mathrm{B}}T\right) B_{12}/B_{21} - 1} = \rho(\varepsilon) = u(\varepsilon, T) = \frac{\varepsilon^3}{\pi^2\hbar^3 c^3}\,\frac{1}{\exp\left(\varepsilon/k_{\mathrm{B}}T\right) - 1}. \tag{4.124}$$

The comparison of the two explicit expressions, which must coincide for *any* temperature, requires the following identities for the coefficients:

$$\frac{B_{12}}{B_{21}} = 1, \qquad \frac{A_{21}}{B_{21}} = \frac{\varepsilon^3}{\pi^2\hbar^3 c^3}, \tag{4.125}$$

known as *Einstein relations*. Relations (4.125), which can be reformulated as

$$B_{12} = B_{21} = \frac{\pi^2 \hbar^3 c^3}{\varepsilon^3} A_{21}, \tag{4.126}$$

were derived under the assumption of equilibrium. Since the coefficients B_{12} B_{21} A_{21} depend only on electromechanical properties of the particles, these relations hold unchanged for arbitrary field conditions. Once one of these coefficients is evaluated, either experimentally or, e.g., by $A_{21} = \gamma_{21}$ of Eq. (2.45), the two others are fixed by Eq. (4.126).

The first equality expresses the symmetric role of the initial and final states in QM, implying that a radiation field induces equal rates of excitation $|1\rangle \to |2\rangle$ and of stimulated emission $|2\rangle \to |1\rangle$. For a strong radiation intensity, such that $B_{21}\, \rho(\varepsilon) \gg A_{21}$, the spontaneous emission rate becomes negligible, and $B_{12} = B_{21}$ implies in particular that $R_{1\to 2} \simeq R_{2\to 1}$, thus rapidly also $n_1 \simeq n_2$ (*saturated transition*). This result clarifies the role of the intense pump beam in the experiment of Fig. 2.11: it saturates the individual components of the $|n = 2\rangle \longleftrightarrow |n = 3\rangle$ transition of hydrogen.

The second relation implies that, for a constant spectral energy density $\rho(\varepsilon)$ (independent of ε), the ratio of spontaneous emission to stimulated emission varies with ε^3. Accordingly, in a low-energy (microwave, IR) transition, stimulated emission tends to prevail, while at higher energy (UV, X-rays) spontaneous emission usually dominates. In contrast, when the radiation is at equilibrium, the ratio of spontaneous to stimulated emission

$$\frac{A_{21}}{B_{21} u(\varepsilon, T)} = \exp\left(\frac{\varepsilon}{k_B T}\right) - 1 \tag{4.127}$$

indicates, unsurprisingly, that thermal radiation is effective in stimulating emission for high temperature $k_B T \gtrsim \varepsilon$ only. Finally, in the context of spectroscopy, the equality $B_{12} = \pi^2 \hbar^3 c^3\, \varepsilon^{-3} A_{21}$ proves that an emission-forbidden transitions is also absorption-forbidden, and, *vice versa*, that a fast allowed transition in emission is also intense in absorption.

4.4.1 The Laser

Ordinary media attenuate traversing electromagnetic waves. Absorption spectroscopies (Fig. 1.3) measure precisely this attenuation. Consider instead an optical medium composed of non-interacting quantum systems which *amplifies*—rather than attenuate—light. For this to occur, the total emission rate needs to exceed absorption: $n_2 R_{2\to 1} > n_1 R_{1\to 2}$, i.e. the ratio

$$\frac{\text{rate of emission}}{\text{rate of absorption}} = \frac{n_2\, [A_{21} + B_{21} \rho(\varepsilon)]}{n_1 B_{12} \rho(\varepsilon)} = \frac{n_2}{n_1} \left[1 + \frac{A_{21}}{B_{21} \rho(\varepsilon)}\right] \tag{4.128}$$

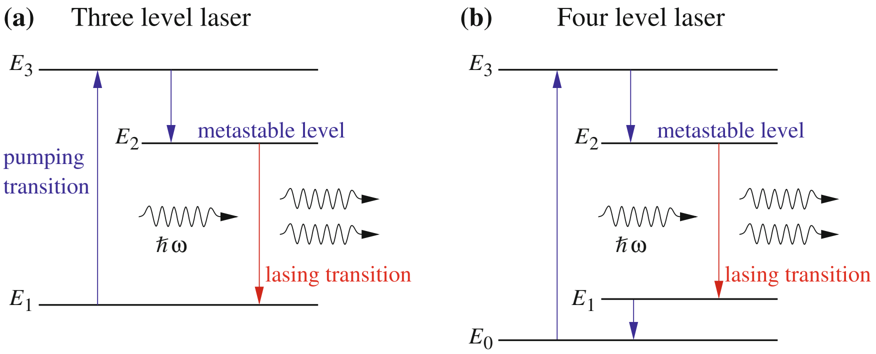


Fig. 4.19 The principle of operation of three- and four-level lasers. These level schemes are two tricks for implementing the population inversion needed to obtain an "active medium", with a larger population of level $|2\rangle$ than that of level $|1\rangle$. The lasing transition $|2\rangle \to |1\rangle$ must occur at a far slower rate than all other indicated *downward* transitions. The latter are fast electric-dipole—allowed transitions

needs to exceed unity. Since the term in brackets approaches unity as soon as a sufficient radiation intensity builds up, this equation tells us that light amplification (emission overcoming absorption) requires a single condition: a population ratio $n_2/n_1 > 1$, i.e. inverted compared to a regular equilibrium population $[n_2]/[n_1] < 1$. An inverted population is a radical deviation from equilibrium, and is thus highly unstable: precisely the prevalence of emission over absorption leads the ensemble of quantum systems spontaneously toward the regular Boltzmann population $[n_2] < [n_1]$. In practice, some kind of electronic or optical *pumping* is needed to produce and sustain the population inversion for an extended period of time, and replace the radiated energy at the expense of an external power source, as in the schemes of Fig. 4.19.

Once an optical *active medium* which amplifies light is realized, light can be channeled through it in a precise direction, by building a resonating one-dimensional (1D) optical cavity around it. For an historically relevant example, see the ruby laser illustrated in Fig. 4.20. A crucial feature of stimulated emission is coherence: the emitted photon is not emitted at random as in spontaneous emission, but prevalently in the same direction and with the same phase and polarization as the stimulating photon, which thus is "cloned".

The use of a long cavity lets photons emitted at odd directions escape basically unamplified, while photons directed along the cavity axis bounce back and forth several times through the active medium, thus getting strongly amplified. As a result, a powerful highly coherent beam of radiation builds up for as long as the population inversion is maintained. A device such as described here, producing a coherent beam of photons by means of Light Amplification by the Stimulated Emission of Radiation is named *laser*.

Diverse commercial applications of such coherent beams extend from the industrial to the consumer side, including telecommunications, optical data storage,

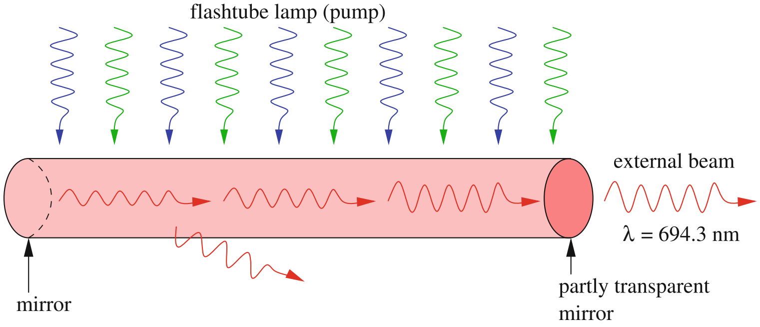


Fig. 4.20 The ruby laser is a three-level solid-state laser. Chrome impurities in an (otherwise transparent) Al_2O_3 crystal rod act as noninteracting quantum systems carrying the levels involved in the population inversion, as in Fig. 4.19a. A *white* flashtube lamp pumps this "active medium" optically, raising many Cr impurities from their ground state $|1\rangle$ to a band of broad short-lived states $|3\rangle$. The latter decay rapidly to metastable state $|2\rangle$. As a further spontaneous decay is very slow, the $|2\rangle \rightarrow |1\rangle$ transition is mainly induced by stimulated emission. Since the energy of state $|2\rangle$ is defined sharply, the emitted radiation is highly monochromatic ($\lambda = 694.3 \pm 0.5\,\text{nm}$). Off-axis photons arising from spontaneous decay escape from the sides of the rod. Repeatedly reflected axially moving photons get amplified and stimulate further coherent emission. The output beam of photons escapes through the partly transparent mirror at one end of the rod

telemetry, cutting, welding, surgery, etc. Lasers play also a fundamental role as irreplaceable tools for research in spectroscopy, photochemistry, ultracold trapped-gas cooling, microscopy, adaptive optics of telescopes, interferometry, etc.

4.5 Final Remarks

The present chapter connects thermodynamics with the principles of equilibrium statistics, based on a simple "democratic" assumption of equal probability of all states of a given energy. We focus on applications to ordinary matter, especially on the few systems (mainly gas phases) where ideal noninteracting systems provide semi-quantitative results.

Eventually, we obtain several quite different formulations and results, which are applicable in different contexts. Table 4.1 summarizes the main formulas and their range of applicability.

Statistical physics becomes more complicated and exciting when moving beyond ideal systems and addressing applications to real interacting systems [29, 30]. Physicists have devised all sorts of techniques to investigate interacting systems both theoretically (mean-field and many-body methods, the renormalization group,…) and experimentally (the definition and measurement of position/momentum/spin correlation functions, high-pressure techniques to investigate phase diagrams in extreme regimes, analysis of fluctuations, finite-size effects, exotic magnetic phases…).

Table 4.1 A summary of the main results of this chapter, with their appropriate applicability range

Physical context	Main relations $[\beta = (k_B T)^{-1}]$	Applicability
General Boltzmann statistics	$Z = \sum_{m'}^{\text{global states}} e^{-\beta E_{m'}}$ $P_m = e^{-\beta E_m}/Z$	Any system at equilibrium
Noninteracting "distinguishable" particles (any T)	$Z = \prod_i Z_i$ $Z_i = \sum_\alpha^{\text{states of particle } i} e^{-\beta E^i_\alpha}$ $P^{(i)}_\alpha = e^{-\beta E^i_\alpha}/Z_i$	Localized or internal degrees of freedom
Noninteracting identical particles (high T)	$Z = (Z_1)^N/N!$ $Z_1 = Z_{1\,\text{tr}} = V/\Lambda^3, \quad \Lambda = \hbar\left(\frac{2\pi}{Mk_B T}\right)^{1/2}$ $PV = Nk_B T$	Ideal-gas translational degrees of freedom
Noninteracting identical bosons (any T)	$Q = \prod_\alpha \left[1 - e^{\beta(\mu - \mathcal{E}_\alpha)}\right]^{-1}$ $[n_\alpha]_B = 1/(e^{\beta(\mathcal{E}_\alpha - \mu)} - 1)$ $PV = -k_B T \sum_\alpha \ln\left(1 - e^{\beta(\mu - \mathcal{E}_\alpha)}\right) = \frac{2}{3}U$	Ideal-gas translational degrees of freedom
Noninteracting identical fermions (any T)	$Q = \prod_\alpha \left[1 + e^{\beta(\mu - \mathcal{E}_\alpha)}\right]$ $[n_\alpha]_F = 1/(e^{\beta(\mathcal{E}_\alpha - \mu)} + 1)$ $PV = k_B T \sum_\alpha \ln\left(1 + e^{\beta(\mu - \mathcal{E}_\alpha)}\right) = \frac{2}{3}U$	Ideal-gas translational degrees of freedom

The present introduction to statistical thermodynamics focuses on equilibrium, and omits *dynamical* out-of-equilibrium quantities. These quantities are needed to discuss transport (e.g. in the next chapter we will define electrical and thermal conductivities in solids) and general hydrodynamic properties derived from the microscopic interactions.

Problems

A ⋆ marks advanced problems.

4.1 ⋆ A gas-phase HCl sample is traversed by a microwave field resonating with the rotational transition between the rotational states of angular momentum $l = 1$ and $l = 2$. Evaluate the ratio between the spontaneous and stimulated emission rates, given the equilibrium distance $R_M = 127\,\text{pm}$ between the proton and the ^{35}Cl nucleus, and the microwave spectral density at resonance $\rho_\nu = 0.250\,\text{J}\,\text{m}^{-3}\,\text{Hz}^{-1}$.

4.2 ⋆ The inside density of "white dwarf" stars can reach approximately $10^{11}\,\text{kg}\,\text{m}^{-3}$. Assume for simplicity that

- these stars are composed by non-interacting protons and electrons in equal number and uniform density;
- that the electrons are ultra-relativistic, with energy $E = c|\mathbf{p}| = \hbar c|\mathbf{k}|$;
- that temperature is 0 K.

Within the above hypothesis, evaluate the pressure of this electron gas.

4.3 To a fair approximation, the valence electrons of metal lithium form a Fermi gas, while positive ions remain essentially immobile at crystalline equilibrium positions. The density of this this material is 535 $\mathrm{kg\,m^{-3}}$. Evaluate the pressure of the electron gas at temperature 0 K. Instead, if one could generate a weakly-interacting *atomic*-^{6}Li gas with the same density, evaluate the pressure of such Fermi gas at temperature 0 K.

4.4 Consider the dissociated single-atom fraction of a fluorine gas at temperature 5,000 K, thus neglecting all diatomic molecules. Neglect also all atoms in electronic states with a configuration other than $1s^2 2s^2 2p^5$. Evaluate the fraction of atoms in the state with total angular momentum $J = \frac{1}{2}$, given that its excitation energy is 50.1 meV above the atomic ground state. Compute also the translational and electronic contribution of each atom to the molar heat capacity at the assigned temperature.

4.5 Comparing gas-phase samples of Cl atomic at equilibrium at temperatures 1,000 and 2,000 K, evaluate the intensity ratio $I(2{,}000\,\mathrm{K})/I(1{,}000\,\mathrm{K})$ of the least energetic absorption line exhibited by a sample in the $3s^2 3p^5 (^2P) \to 3s^2 3p^4 4s (^2P)$ transition, due to the different thermal population of the ground and first-excited state of the $3s^2 3p^5 (^2P)$ configuration. These $3s^2 3p^5 (^2P)$ states are 109.4 meV apart.

4.6 The three atomic levels associated to the ground configuration $3d^2 4s^2\,{}^3F$ of titanium are found at energy 0, 0.02109, and 0.04797 eV. Evaluate the contribution of these electronic excitations to the molar heat capacity of a vapor of monoatomic titanium at temperature 1,000 K, and their fractional contribution to the total heat capacity of the gas.

4.7 An insulating solid contains non-interacting atomic impurities (at a number-density level $3 \times 10^{21}\,\mathrm{m^{-3}}$) characterized by localized levels F (threefold degenerate) and A (nondegenerate), separated by an energy $E_A - E_F = 50\,\mathrm{meV}$. Evaluate the contribution of these impurities to the heat capacity per kilogram of this solid at $T = 700\,\mathrm{K}$, given that the material's density is 2,300 $\mathrm{kg/m^3}$.

4.8 Evaluate the ratio of the heat capacity per unit volume of the electromagnetic fields in a blackbody cavity at temperatures 10,000 K and 4,000 K. Evaluate also the ratio between the total radiated power at the two indicated temperatures.

4.9 Evaluate the mean speed $\langle |v| \rangle$ and the pressure of the conduction electrons in the ground state of metallic gold (density $19.3 \times 10^3\,\mathrm{kg\,m^{-3}}$, one electron per atom in the conduction band), in the approximation of free non-interacting electrons.

4.10 * Approximate the spectrum emitted by a furnace with blackbody radiation. The radiated power (in $\mathrm{W\,m^{-2}}$) in the wavelength range between 3,150 and 3,250 nm (infrared) is measured. One observes that when the furnace absolute temperature is doubled, this power increases by a factor 10. Evaluate the final furnace temperature. At such temperature, what is the ratio of the radiated power in the range 3,150–3,250 nm to the radiated power in the visible range 695–705 nm? [Suggestion: evaluate integration approximately as finite sums, taking advantage of the narrowness of the integration intervals.]

4.11 Estimate, (with at least 10 % precision) the temperature needed to produce with a 0.1 % efficiency, soft X-rays of energy $\geq$ 10 eV using a thermal blackbody source. [Suggestions: (a) when $x \gg 1$, the error in the approximation

$(e^x \pm 1)^{-1} \simeq e^{-x}$ is negligible; (b) $\int_0^\infty x^3/(e^x - 1)dx = \pi^4/15$; (c) approximate solutions of non-algebraic equations can be obtained by numeric bisection.]

4.12 Approximate a human body to a blackbody at temperature 37 °C. Evaluate the total power that it irradiates, assuming a surface area $1.8\,\mathrm{m}^2$. Taking also into account the power that reaches the human body when inside a blackbody cavity, at what temperature should this cavity be placed for the net power lost by the human body to be reduced by a factor 10?
[Recall the expression for the Stefan-Boltzmann constant: $\pi^2 k_B^4/(60\,\hbar^3 c^2)$.]

4.13 Consider a metallic sodium sample (density $\rho = 950\,\mathrm{kg\,m^{-3}}$). Assuming the free non-interacting Fermi model for the conduction electrons, evaluate the electronic pressure. Assume that 10 % of the sodium atoms is replaced by aluminum atoms, with no change in the crystal structure (even the lattice spacing remains unchanged). Evaluate the variation of the electronic pressure compared to pure sodium due to the presence of 3 (rather than 1) electrons per aluminum atom in the conduction band.

4.14 * The maximum phonon frequency of NaCl is $\nu_{\max} = 5\,\mathrm{THz}$. Assume that all vibrational modes whose frequency is smaller than of equal to $\nu_{\max}$ contribute to electromagnetic radiation absorption, to the extent that NaCl is equivalent to a black body in the $0 - \nu_{\max}$ frequency interval. Assume moreover that at frequencies larger than $\nu_{\max}$ the solid neither absorbs nor emits electromagnetic radiation. Estimate (with less than 10 % error) the total radiated power emitted by a NaCl crystal with a $1\,\mathrm{cm}^2$ surface kept at temperature 2,000 K.
[Recall the density of states of the electromagnetic fields in a cavity of volume V: $g(\varepsilon) = V\varepsilon^2/(\pi^2\hbar^3 c^3)$.]

Chapter 5
Solids

Macroscopic systems realize a thermodynamic equilibrium state in a balance between the tendency of internal energy to decrease and that of entropy to increase. Temperature tunes the balance in favor of the entropic contribution over the energetic one: entropy prevails at high temperature, energy at low temperature. The entropy of the solid state, where each atom sits most of the time around a definite position, is usually smaller than in fluid states. Indeed, experience shows that most materials solidify as temperature is lowered sufficiently.

The internal energy consists of a kinetic plus a potential term. The translationally invariant kinetic energy T_n tends to favor states characterized by delocalized and uncorrelated positions of individual atoms, not unlike entropy. On the contrary, the adiabatic potential energy V_{ad} takes advantage of characteristic optimal interatomic separations (see e.g. Fig. 3.1) and angles, and attempts therefore to impose strong positional localization and correlations. The solid state signals the prevalence of the adiabatic potential energy over the atomic kinetic energy in Eq. (3.9). For almost all materials at zero temperature, V_{ad} prevails over T_n, thus leading to solid states. The one remarkable exception is helium, which at ordinary pressure remains fluid (a *quantum fluid*) down to zero temperature, due to its exceptionally weak interatomic attraction associated to a scarce atomic polarizability (see Table 3.1). As atoms are squeezed together by an applied pressure, the repulsive part of the He-He interaction starts to dominate over T_n, and even helium acquires low-temperature solid phases. These phases are *quantum solids* because zero-point motion, tunneling, and all sorts of quantum-kinetic effects play a significant role.

The tendency of atoms in a solid to stick together is measured by the "total binding energy" (in molecular language, the energy necessary to disaggregate the compound into its *atomic* components). In the context of solids, this quantity is referred to as the lattice *cohesive energy*, or lattice energy, and it is often expressed per atom, or per formula unit, or per mole. Like for molecules, both the V_{ad} and T_n contributions vanish in the "atomized" state (assuming $V_{\text{ad}} = 0$ in the atomized state), whereas in the solid state, $[V_{\text{ad}}] < 0$. In the solid state the kinetic term $[T_n] \geq 0$ is much smaller than $|[V_{\text{ad}}]|$, see Fig. 3.15.

N. Manini, *Introduction to the Physics of Matter*,
Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-14382-8_5

Solid matter is characterized by long-distance rigidity: a force applied to one or few of the atoms in a solid sample acts through the whole sample, which accelerates maintaining its average shape unchanged. This is due to the ability of solids (as opposed to fluids) to resist shear forces. This basic macroscopic property, on which our daily experience relies, is far from trivial from the point of view of the microscopic equations governing the dynamics of electrons and nuclei composing a solid. Indeed, our analysis of the adiabatic potential acting directly between two atoms at large distance reveals that V_{ad} always decays with a rather fast power law R^{-6}. As the number of atoms at distance R from a given atom grows as R^2, the total interaction energy with faraway atoms decays as R^{-4}. In practice, each atom interacts significantly only with the atoms sitting in its close neighborhood, of a few nm^3 say. Long-distance rigidity therefore cannot be related to long-range forces. This means that, in a solid, short-range forces propagate from one atom to the next ones, and from those to farther atoms again and again, through the whole sample.

5.1 The Microscopic Structure of Solids

Many solids exhibit ordered microscopical structures, but highly disordered solids are very frequent as well. Before coming to the experimental evidence for the ordered structure of many solids, we try to understand why regular spatial arrangements of atoms can and do emerge spontaneously.

In our initial study of many-atoms system (Chap. 3), we analyzed the typical shape of the adiabatic potential (Fig. 3.1) for a diatom. We also found that the adiabatic potential of many atoms is an explicit function of the relative positions of all of them, including all distances and angles. For exceptionally simple systems, such as the noble-gas elements, the total adiabatic potential of N_n atoms is fairly well approximated by a sum of 2-body terms (e.g. of the Lennard-Jones type)

$$V_{\mathrm{ad}}\left(R_1, R_2, \ldots, R_{N_n}\right) = \sum_{\alpha<\alpha'}^{N_n} V_2(|R_\alpha - R_{\alpha'}|). \tag{5.1}$$

Let us neglect the nuclear kinetic energy: the state of minimum energy of two such atoms is realized by placing them at the equilibrium distance R_{M} of the potential V_2, with an adiabatic energy equal to the depth $-\varepsilon$ of the potential well.[1] To start, assume that all equal atoms are constrained to move along a line: a third added atom can join the two others on either side, at approximately a distance R_{M} from the nearest atom, as illustrated in Fig. 5.1. Neglecting the weak attraction of second- third- etc. -neighbors (see Fig. 5.1), the adiabatic potential energy decreases to approximately -2ε. Likewise, N_n atoms along a chain place themselves at almost perfectly regular

---

[1] For the Lennard-Jones potential—Eq. (3.19)—$R_{\mathrm{M}} = \sqrt[6]{2}\,\sigma$, with $V_2(R_{\mathrm{M}}) = V_{\mathrm{LJ}}(R_{\mathrm{M}}) = -\varepsilon$.

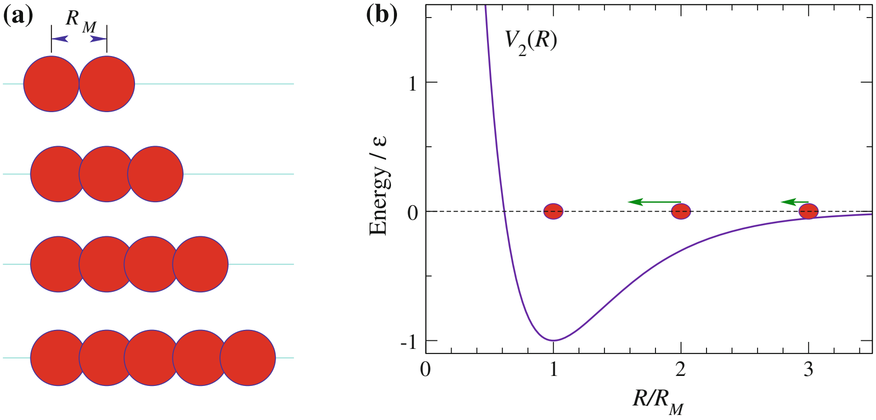


Fig. 5.1 **a** A 1D solid constructed by the successive addition of atoms. The atoms after the second add at fairly regular distances, slightly smaller than R_M, the optimal equilibrium distance of the 2-body adiabatic potential. **b** The reason of the slightly smaller equilibrium separation in the solid than in the diatom: if the separation was R_M, forces to second, third,… neighbors would all be attractive and uncompensated, thus some energy is gained by contracting the interatomic separation. This contraction is tiny, as V_2 explodes at short distance, and even tinier at the surface than in the bulk, due to local deficiency of second, and farther neighbors

distances,[2] and gain a total cohesive energy $\simeq(N_n - 1)\varepsilon$, i.e. essentially ε per atom. Depending on the precise shape of V_2, the actual cohesive energy turns out slightly larger due to the attraction of second and farther neighbors.

In 1D, regularity is trivial and indeed unavoidable for atoms of a single species. In more than 1D, the greater geometric freedom allows several regular and irregular atomic patterns to form. In 2D, a third atom can join two to form an equilateral triangle; further atoms that join the cluster find a lowest-energy arrangement by progressively building a triangular lattice (also called hexagonal lattice), as illustrated in Fig. 5.2. Each atom (except the few ones at the border) is surrounded by 6 nearest neighbors: it thus forms 6 bonds, each shared by two atoms. Accordingly, the cohesive energy per atom is approximately 3ε. The attraction of farther neighbors adds a small correction to the total cohesive energy. The important message here is that two-body interactions favor configurations of *maximal coordination*, i.e. arrangements where each atom is surrounded by as many nearest neighbors as geometry allows.[3]

---

[2] Slight distortions occur at the surface. Deep inside the bulk, however, each atom is subject to basically the same interactions as its neighboring ones, thus it reaches an equilibrium position relative to the surrounding ones which involves perfectly regular spacings.

[3] For comparison, if the atoms arranged themselves in a square lattice, each atom would bind to 4 nearest neighbors, rather than 6. The resulting cohesive energy per atom would amount to $\simeq 2\varepsilon$ instead of $\simeq 3\varepsilon$: the macroscopic cohesive energy difference $\Delta U = U^\mathrm{square} - U^\mathrm{triang} \simeq N_n\varepsilon$ indicates that the square lattice is strongly unstable and deforms spontaneously to the triangular lattice shape.

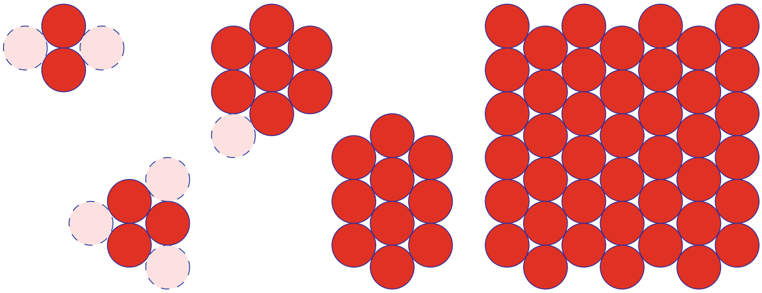


Fig. 5.2 A 2D *solid* constructed by successive addition of atoms. The basic unit is the equilateral *triangle*, which repeats itself indefinitely in space, so that each atom in the bulk is (maximally) coordinated to 6 other atoms. Contraction due to 2nd-, 3rd-,…-neighbor attraction is ignored in this figure

The principle of maximal coordination has important consequences also for 3D solids.[4] 4 atoms maximize their coordination by placing themselves at the vertexes of a regular tetrahedron. Extra atoms extend this basic unit in space following either of the following regular patterns: the *face-centered cubic* (fcc) lattice and the *hexagonal close-packed* (hcp) lattice. As illustrated in Fig. 5.3, both lattices are the result of a regular stacking of 2D triangular lattices. In both hcp and fcc lattices, the second layer is stacked above the first one so that the atoms sit on top of the centers of half the triangles of the lower layer. The third layer is also stacked on top of half the centers of the second-layer triangles: in the hcp directly above the atoms of the first layer, while in the fcc above the remaining triangles. In both patterns, each atom is surrounded by 12 nearest neighbors, thus the cohesive energy is approximately 6ε per atom. When the attraction of farther neighbors is accounted for, the cohesive energy is usually marginally more favorable to fcc than to hcp.

At low temperature the noble gases are indeed observed to crystallize in fcc structures. The optimal equilibrium distance (accounting for all farther-neighbors interactions) for the Lennard-Jones fcc solid equals $0.971\,R_\mathrm{M} = 1.09\,\sigma$, with a total cohesive energy per atom of 8.6ε (significantly more bound than the 6ε nearest-neighbor estimate). By plugging the parameters of Table 3.1 in this simple model, one obtains the bond lengths and cohesive energies of Table 5.1. Not unexpectedly, the experimental energy is generally slightly larger (less cohesive) than the prediction of the simple Lennard-Jones model, where the ionic kinetic energy T_n and the associated zero-point motion are neglected. The good overall agreement confirms the concept of maximizing the coordination, leading to compact structures similar to the packing of hard spheres, typically fcc, a concept valid whenever the adiabatic potential can be decomposed into 2-body terms, as in Eq. (5.1).

The noble gases (and mixtures thereof) are not the only system where the approximation of a 2-body adiabatic potential works: it can describe successfully the overall structure of many solids formed by “spherically symmetric” close-shell molecules,

---

[4] 3D structures are visualized better in 3D than with flat projections. The atomic coordinates depicted in Fig. 5.3 and other figures below are available for download and 3D visualization at the web site of Ref. [25].

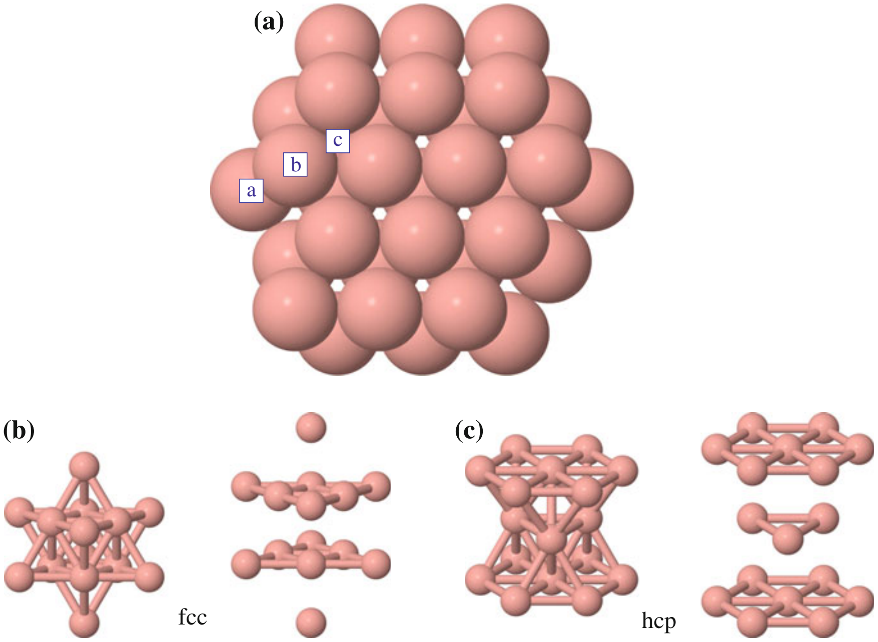


Fig. 5.3 A close packing of spheres (e.g. cannonballs) in 3D is obtained by stacking close-packed 2D arrays, namely triangular lattices. If in panel **(a)** a third layer is superposed on sites of type "c", the fcc stacking is initiated. If instead the third layer is stacked on top of the spheres of layer "a", the hcp crystal is realized. Panel **(b)** illustrates the conventional fcc cell, realized as a stacking sequence of type abcabc…Panel **(c)** illustrates the same concept for hcp: an ababab…stacking

Table 5.1 The estimate of the equilibrium nearest-neighbor interatomic separation a and cohesive energy per atom for solid noble-gas elements based on the fcc Lennard-Jones model and parameters (Table 3.1), compared to experimental determinations [10]

Element	a (pm)	$1.09\,\sigma$ (pm)	$\|U\|/N_n$ (meV)	8.6ε (meV)	T_{melt} (K)
	Exp	Theory	Exp	Theory	Exp
Ne	313	300	20	27	24.6
Ar	375	371	80	90	83.8
Kr	399	401	110	124	115.8
Xe	433	443	170	167	161.4

Discrepancies of the order of few percent are not surprising, especially for the light Ne, in view of the neglect of the kinetic energy of the nuclei. Note the strong correlation between the cohesive energy per particle and the melting temperature T_{melt} of the solid ($|U|/N_n \simeq 12\,k_B T_{\mathrm{melt}}$)

e.g. methane CH_4 [35]. With suitable modifications, similar 2-body models can describe other *molecular solids* (e.g. H_2, N_2, Cl_2) where again weak van der Waals interactions provide cohesion, but the asymmetry of the individual molecules can favor different lattice structures.

The hitherto discussed molecular solids, where each molecular unit retains many of its molecular properties and is only weakly bound to other units, constitute a marginal class of solids, by no means the most typical one, like the Ar_2 dimer is a rather exotic example of a molecule. Contrary to molecular solids, electrons of the outer atomic shells modify substantially their quantum state when they belong to a *covalent* or *metallic solid*, in pretty much the same way that the electronic state of H N and O is modified in forming the covalent molecules H_2, N_2, H_2O, as discussed in Sect. 3.2. A covalent or metallic solid can indeed be viewed as a huge molecule, where the molecular bonding extends to the whole sample. Experimentally, cohesive energies per atom $|U|/N_n$ of solids are comparable to the bond energies of covalently bound diatomic molecules, i.e. of the order of several eV. These are of course much larger than those of noble-gas solids (Table 5.1); for example we report the cohesive energies per atom of a few elemental solids: Li 1.65 eV, C (diamond) 7.36 eV, Si 4.64 eV, Fe 4.29 eV; Cs 0.83 eV. *Ionic crystals* (e.g. NaCl) exhibit similar cohesive energies, of several eV per ion pair.

To estimate quantitatively the cohesive energy of a covalent or metallic solid, it is necessary to study the dynamics of its electrons in detail. Like for many-electron molecules, in practice, reliable estimates can be computed by means of detailed self-consistent calculations. Before coming back to the electronic states of solids in Sect. 5.2, observe that it is to be expected that the adiabatic potential associated to such nontrivial electronic states depends strongly on all bond lengths and angles, pretty much as it does in molecules where it enforces relatively rigid equilibrium molecular geometries (Fig. 3.8). Therefore, in covalent and metallic solids, the 2-body approximation Eq. (5.1) is bound to fail completely, and different structures, other than fcc, are to be expected, depending on the detailed chemistry and thus on the relevant V_{ad} involved. Indeed, elemental solids show several different ordered structures including (beside fcc): hcp, *body-centered cubic* (bcc), *diamond*, and others. Long-range crystalline order reaches spectacular levels of perfection, for example in industrial-grade Si single crystals, where a sub-nanometer unit cell repeats itself over and over in three dimensions for distances exceeding 1 m (Fig. 5.4). We shall soon discuss these structures within the standard formalism of periodic crystals, based on the infinite periodic repetition in space of a small piece of matter.

A perfect infinitely repeated lattice is an idealization which is never exactly realized in nature. In a real material, the crystalline structure is doomed to contain *defects*. A localized defect, such as a vacancy or an interstitial atom (Fig. 5.5) raises the total energy by a few times the typical bond energy ε of an atom to its neighbors, and this has an irrelevant effect on the total cohesive energy per atom U/N_n, in the large-N_n limit. At equilibrium at low temperature, Boltzmann statistics predicts a small relative concentration (of the order of $\sim e^{-\beta\varepsilon}$) of localized defects to survive. The modest difference in energy between the fcc and hcp lattices makes the creation even of extended defects (e.g. dislocations—Fig. 5.6a, or stacking faults—Fig. 5.7) likely. The energy cost of extended defects is macroscopically large and should therefore suppress them strongly at equilibrium. However defects remain easily "frozen" within the solid, the typical time for a defect to drift out of a macroscopically large sample being often astronomically large. As a result, at low temperature a solid often remains locked

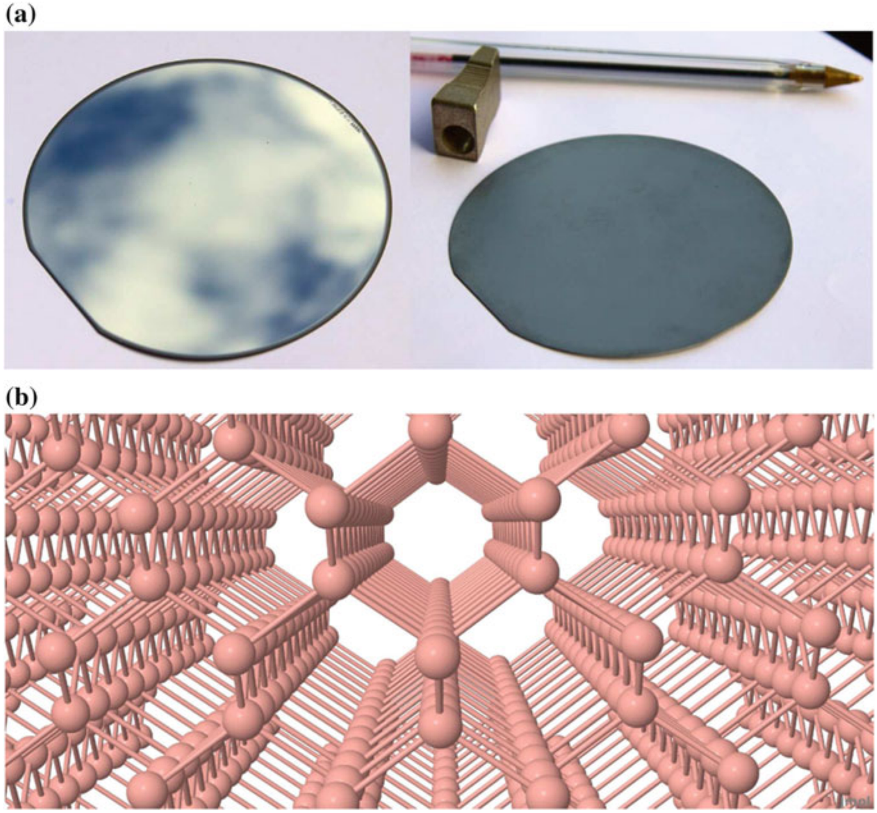


Fig. 5.4 **a** Two 0.3 mm thick wafers sliced from ~1 m long silicon boules, i.e. single-crystal ingots. The diameters of these wafers are 100 mm (*left*) and 50 mm (*right*)—larger ones are available commercially. The surface of the wafer at the *left* is polished, thus mirror-reflecting, while the picture the *right* focuses on the unpolished side. (The wafers were kindly provided by A. Podestà and R. Manenti.) **b** A *balls-and-sticks view* "from the inside" of the silicon crystal lattice structure down the (110) direction. Each atom binds four other nearest-neighbor atoms only. This same geometric arrangement describes diamond (C) and germanium as well

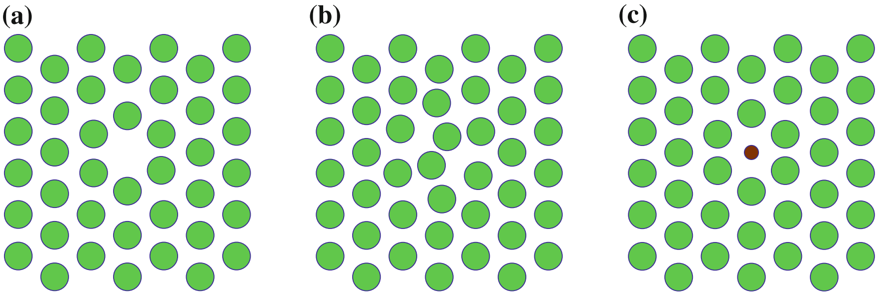


Fig. 5.5 Point defects in an otherwise perfect triangular lattice: **a** a vacancy, **b** an interstitial atom, and **c** an impurity atom

Fig. 5.6 **a** A dislocation in a simple-cubic lattice. **b** The motion of dislocations reduces substantially the shear resistance of a real crystal relative to a defect-free crystal

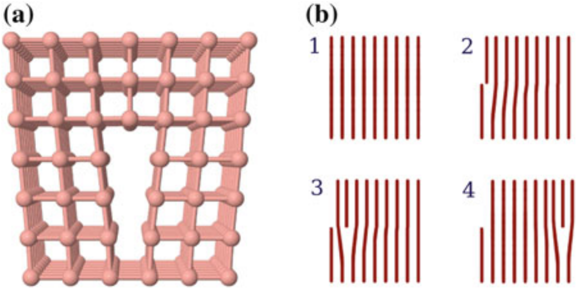


Fig. 5.7 Side view of a stacking fault in a fcc crystal. The defective interface involves an irregular relative stacking of two successive (111) hexagonal planes, see Fig. 5.3

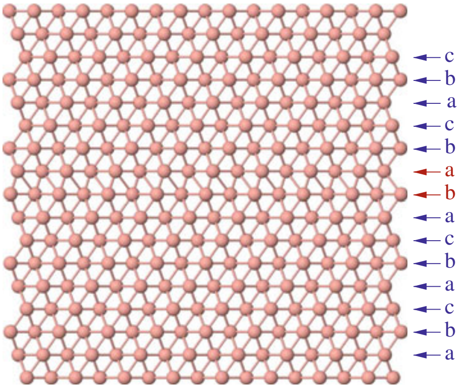
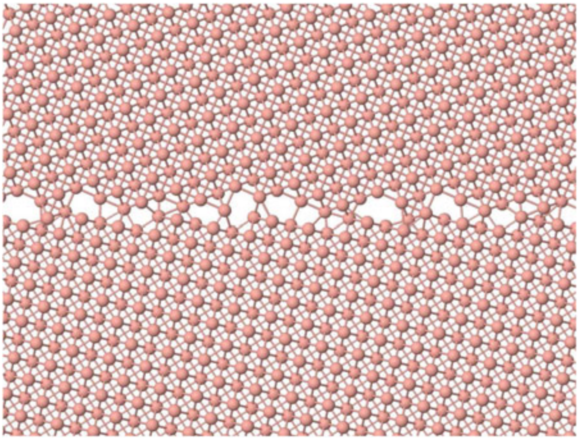


Fig. 5.8 A grain boundary in a fcc crystal viewed *down* the (100) direction



in a metastable non-equilibrium state with a finite (often large) concentration of defects depending on preparation (see the discussion of ortho- and para-hydrogen of Sect. 4.1.1 for a conceptually simpler example of a metastable non-equilibrium state surviving for a long time). Extended defects of the type of Figs. 5.6, 5.7 and 5.8

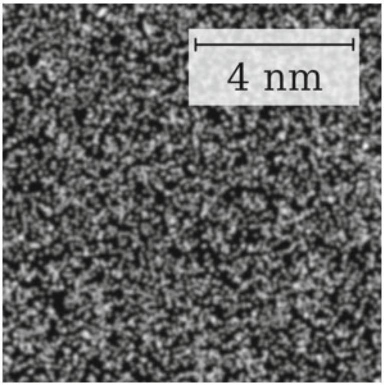


Fig. 5.9 Simulated atomic-resolution microscopy image of an amorphous zirconium (Zr) alloy. Contrasted to, e.g., Fig. 1.1, atoms are arranged randomly, due to the noncrystalline glassy structure of this material

are crucial for understanding the plastic deformations of real crystals under strain. Ultimately, even without internal defects, real solids are never ideal because the lattice periodicity must end at the inevitable terminating surface or interface or grain boundary (see Fig. 5.8).

In many materials (e.g. multiple-component off-stoichiometric compounds, glasses, polymers, alloys,…), the cost of the formation of defects is so small that it is highly nontrivial (often impossible) to obtain crystalline samples. These materials form amorphous solids, where no long-range lattice order is present (Fig. 5.9): their microscopic structure often resembles that of a frozen liquid. The formalism that we are going to set up for crystals is mostly useless for amorphous systems: their investigation requires more advanced tools, exceeding the scope the present basic textbook.

The tendency to grow and to break along flat planes (*cleave*) at fixed characteristic relative angles strongly hints at crystalline order in many solids, e.g. gemstones, see Fig. 5.10. An even more compelling evidence is provided by the diffraction of X-rays, neutrons and electrons of wavelengths in the a_0 region. As Fig. 5.11 illustrates, several

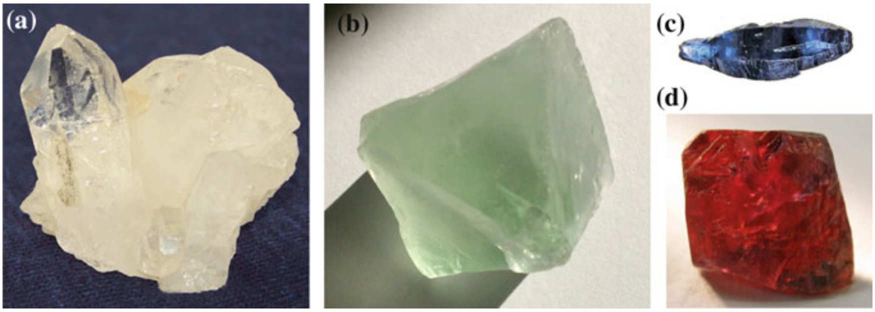


Fig. 5.10 A few naturally grown minerals, resulting typically in aggregates of ~mm-size defective crystals. **a** Quartz (SiO_2). **b** Fluorite (CaF_2). Corundum Al_2O_3: **c** a Ti-doped (*sapphire*) sample from Sri Lanka [Rob Lavinsky, iRocks.com—CC-BY-SA-3.0] and **d** a Cr-doped (*ruby*) sample from Tanzania [Rob Lavinsky, iRocks.com—CC-BY-SA-3.0]

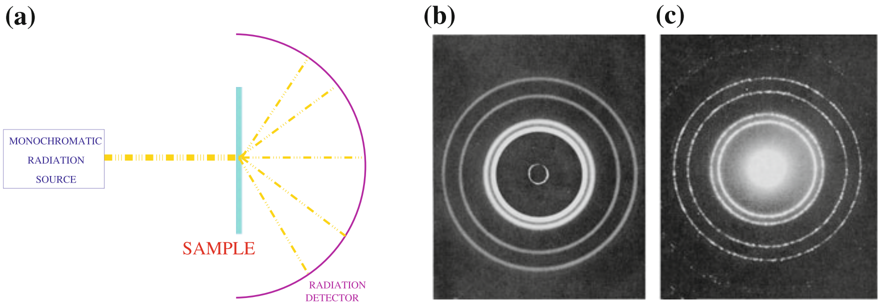


Fig. 5.11 **a** The scheme of a powder or polycrystalline sample diffraction experiment. The "radiation" beam can consist in any wave field interacting with matter, typically X-rays, neutrons, or electrons. The sample must be thin compared to the attenuation length of that radiation in that material. The diffraction patterns made by a beam of **b** X-rays and of **c** electrons passing through the same thin Al foil. The angles characterizing the diffraction rings are functions of the lattice periodicity and the wavelength of the incident radiation

wave probes of wavelength in the correct range interact with crystals and generate diffracted beams, as expected of regular, periodic, arrays of scatterers.[5] We proceed to introduce the basic mathematics describing periodic solids, with the central concepts of direct and reciprocal lattice, and employ the related Fourier analysis to understand diffraction experiments.

5.1.1 Lattices and Crystal Structures

The basic property of a crystalline solid is the essential equivalence of many different regions in space. Any Ne atom in its fcc arrangement "sees" an essentially equivalent environment, unless it lies close to the crystal surface or to some defect. In a sufficiently "clean" crystal, most atoms are far enough, say at least 5 neighbors away, from the nearest imperfection. The fields experienced by one such "bulk" atom equal those the same atom would feel if it belonged in a perfectly periodic structure. It thus makes sense to address many properties of crystalline solids by modeling them as ideal crystals extending through all space.

The main symmetry of a crystal is a *discrete translational symmetry*.[6] Given a point **r** in the crystal, perfectly equal physical properties (including electric potential,

---

[5] Occasional compounds have a quasicrystalline structure: they are perfectly ordered *non-periodic* materials.

[6] In an isolated atom, the rotational symmetry commuting with the effective single-electron Hamiltonian makes the angular momenta l_i of individual electrons good quantum numbers, used to label atomic states such as, e.g., $1s^2 2s^2 2p^4$. Sections 5.1.1 and 5.1.2 address the similar problem to determine and diagonalize the symmetry operators of electrons in a crystal, in order to provide electrons with appropriate quantum numbers.

electric field, mass and charge density, current, etc.) are observed at all other points

$$\mathbf{r}' = \mathbf{r} + \mathbf{R}, \quad \text{with } \mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \tag{5.2}$$

with n_j arbitrary integers, and $\mathbf{a}_j$ three linearly independent vectors. All points of the type $\mathbf{R}$ in Eq. (5.2) form an infinite array extending through space, named a *Bravais lattice*. The vectors $\mathbf{a}_j$ are said to *generate the lattice*. They are called *primitive* if, for any $\mathbf{r}$, the points $\mathbf{r}'$ defined by Eq. (5.2) are all the points which have equal physical properties as $\mathbf{r}$ (that means that $\mathbf{a}_j$ are taken "as short as possible", to make the array of $\mathbf{R}$ points as dense as possible).

As a simplest example, Fig. 5.12 illustrates these ideas for a 2D lattice. Figure 5.13 shows a portion of a simple-cubic crystal, viewed from different angles. Primitive

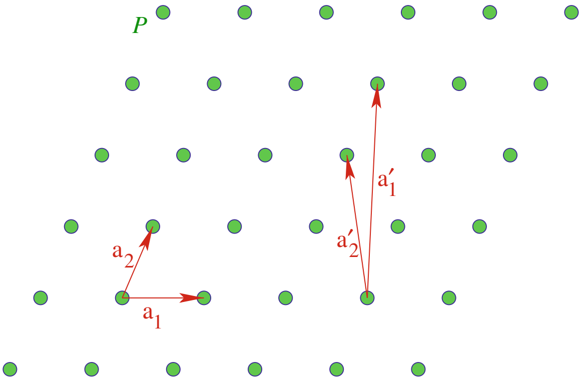


Fig. 5.12 A finite portion of a generic 2D lattice generated by the primitive vectors $\mathbf{a}_1$ and $\mathbf{a}_2$. Other equally good primitive vectors $\mathbf{a}'_1$ and $\mathbf{a}'_2$ are indicated, based on a different origin. Any lattice point $\mathbf{R}$, can be expressed as $n_1\mathbf{a}_1 + n_2\mathbf{a}_2$, for example $P = -1\,\mathbf{a}_1 + 4\,\mathbf{a}_2 = -4\,\mathbf{a}'_1 + 8\,\mathbf{a}'_2$

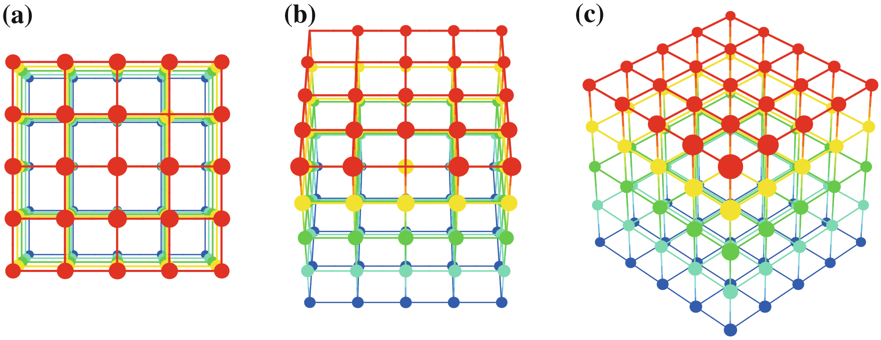


Fig. 5.13 A finite portion ($5 \times 5 \times 5$ lattice spacings) of a simple-cubic lattice observed from **a** the $\hat{\mathbf{z}}$ direction (001), **b** the $\hat{\mathbf{y}} + \hat{\mathbf{z}}$ direction (011), and **c** the $\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}$ direction (111)

vectors can be chosen as $a\hat{\mathbf{x}}$, $a\hat{\mathbf{y}}$, $a\hat{\mathbf{z}}$, i.e. orthogonal and all of the same length a, the side of the smallest cube in the lattice. The drawn portion of the crystal includes 125 points generated by $5 \times 5 \times 5$ consecutive values of the n_j indexes. The only known example of a simple-cubic equilibrium crystal structure (with one atom at each Bravais-lattice point) is polonium.

Other examples of 3D Bravais lattice—already encountered above—are the fcc and bcc lattices sketched in Figs. 5.14 and 5.15. The equilibrium structure of many elements are fcc, e.g. aluminum, nickel, copper, silver, gold, lead, or bcc, e.g. lithium, vanadium, chromium, iron, molybdenum, tungsten. These lattices are built by adding sites to the simple-cubic lattice (in the fcc, at the center of each face of the cubes, in the bcc at the body center of the cubes). The added sites are perfectly equivalent to the original sites at the cube corners. The lattice-point density of fcc is four times and that of bcc is twice that of simple cubic of the same cube side a. For fcc and bcc, the same cubic-lattice vectors $a\hat{\mathbf{x}}$, $a\hat{\mathbf{y}}$, $a\hat{\mathbf{z}}$ as for the simple-cubic lattice are often conveniently used. However, these are not primitive: Fig. 5.16 depicts a standard choice of primitive lattice vectors.

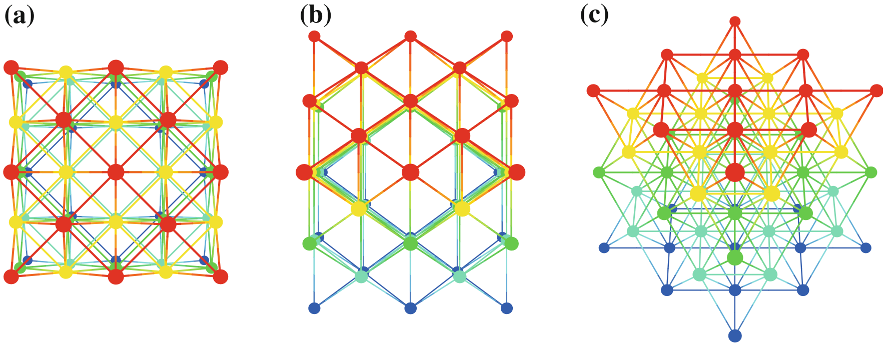


Fig. 5.14 A finite portion ($2 \times 2 \times 2$ lattice spacings) of a face-centered cubic (fcc) lattice observed from **a** the $\hat{\mathbf{z}}$ direction (001), **b** the $\hat{\mathbf{y}} + \hat{\mathbf{z}}$ direction (011), and **c** the $\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}$ direction (111)

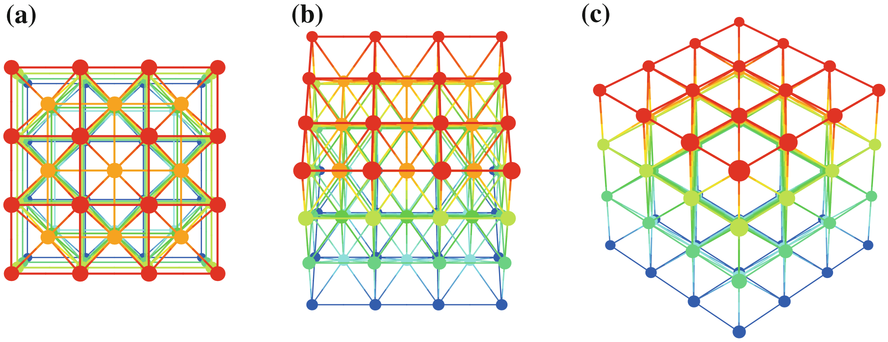


Fig. 5.15 A finite portion ($3 \times 3 \times 3$ cubic lattice spacings) of a body-centered cubic (bcc) lattice observed from **a** the $\hat{\mathbf{z}}$ direction (001), **b** the $\hat{\mathbf{y}}+\hat{\mathbf{z}}$ direction (011), and **c** the $\hat{\mathbf{x}}+\hat{\mathbf{y}}+\hat{\mathbf{z}}$ direction (111)

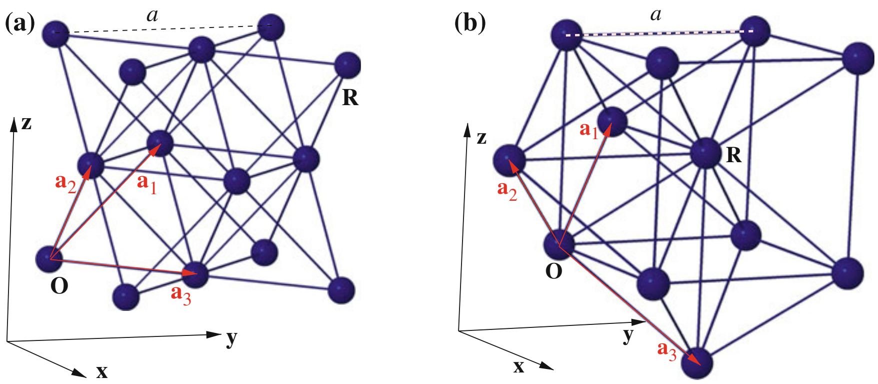


Fig. 5.16 **a** Primitive lattice vectors for the fcc Bravais lattice: $\mathbf{a}_1 = \frac{a}{2}\,(\hat{\mathbf{y}} + \hat{\mathbf{z}})$, $\mathbf{a}_2 = \frac{a}{2}\,(\hat{\mathbf{z}} + \hat{\mathbf{x}})$, $\mathbf{a}_3 = \frac{a}{2}\,(\hat{\mathbf{x}} + \hat{\mathbf{y}})$. **b** Primitive lattice vectors for the bcc lattice: $\mathbf{a}_1 = \frac{a}{2}\,(\hat{\mathbf{y}} + \hat{\mathbf{z}} - \hat{\mathbf{x}})$, $\mathbf{a}_2 = \frac{a}{2}\,(\hat{\mathbf{z}} + \hat{\mathbf{x}} - \hat{\mathbf{y}})$, $\mathbf{a}_3 = \frac{a}{2}\,(\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}})$. As an example, in both lattices the translation marked by **R** can be expressed as $\mathbf{R} = \mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3$

The primitive vectors can be used as three converging edges which define a parallelepiped of volume $V_c = \mathbf{a}_1 \times \mathbf{a}_2 \cdot \mathbf{a}_3$. This parallelepiped contains all "different", translationally-inequivalent points $\mathbf{r}$ in space: any other $\mathbf{r}'$ lying outside this parallelepiped is equivalent to some $\mathbf{r}$ inside, to which it can be translated by using Eq. (5.2) with suitable n_j. This parallelepiped is an example of a *primitive cell*, or *unit cell*: a minimal volume which, by applying lattice translations **R**, fills up the whole space without overlapping. The primitive cell contains all the relevant information about the entire periodic crystal: what happens outside the primitive cell amounts to essentially boring repetitions of what goes on inside. According to definition, a primitive cell needs not be a parallelepiped (in 2D, a parallelogram—see Fig. 5.17).

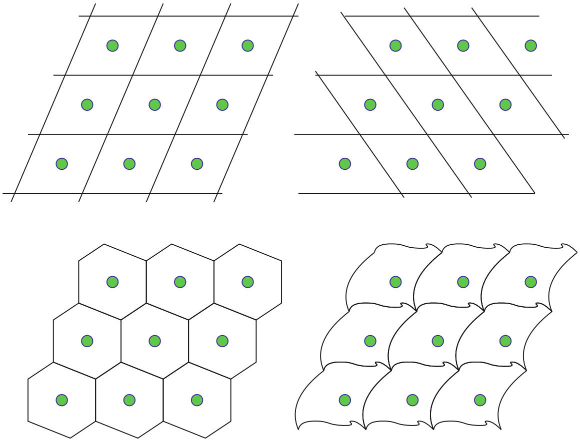


Fig. 5.17 Several possible choices for the primitive cell of the same 2D lattice

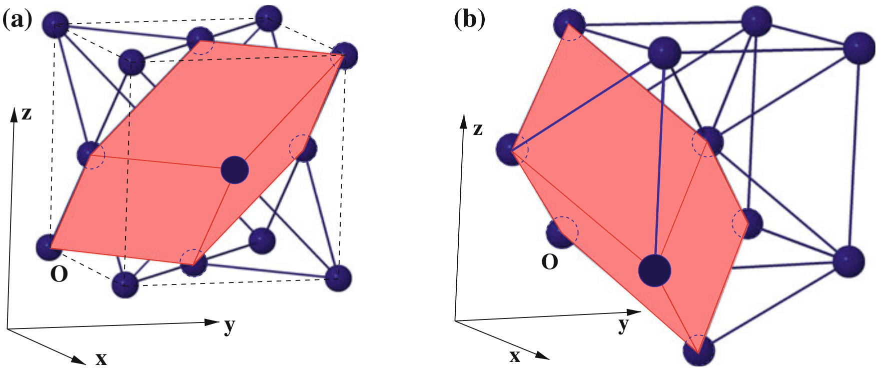


Fig. 5.18 The primitive cells of the fcc (**a**) and bcc (**b**) Bravais lattices. The volume of the fcc primitive cell is 1/4 of the volume a^3 of the conventional cubic cell (*dashed edges*). The volume of the bcc primitive cell is 1/2 of the volume a^3 of the conventional cell

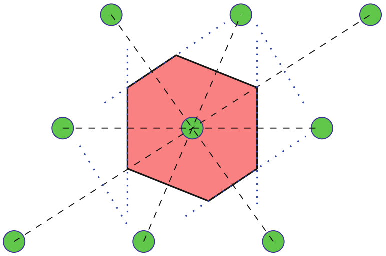


Fig. 5.19 The geometric construction of the Wigner-Seitz primitive cell of a generic 2D lattice. The Wigner-Seitz cells of all 2D lattices are hexagons, except when the lattice is *rectangular* or *square*

Each primitive cell contains, in particular, one and only one Bravais-lattice point, see e.g. Fig. 5.18 for the fcc and bcc lattices.

Given a lattice point $\mathbf{R}$, the *Wigner-Seitz cell* is the set of all points $\mathbf{r}$ closer to $\mathbf{R}$ than to any other lattice point $\mathbf{R}'$ (see for example Fig. 5.19). One can show that the Wigner-Seitz cell is indeed a primitive cell. This special choice of primitive cell retains the full symmetry of the lattice, and does not suffer the arbitrariness (illustrated in Fig. 5.12) of the choice of the primitive vectors. Figure 5.20 shows the Wigner-Seitz cells for two common lattices.

While in many crystals exactly one atom happens to sit in each primitive cell of a Bravais lattice as, e.g., in solid neon or aluminum (both fcc crystals), even more frequently several atoms belong in the same primitive cell, which is then repeated periodically in space. For example, the nuclei of solid NaCl (and many similar compounds such as LiCl, NaBr, KI, AgF, CaO, BaSe,…) sit around simple-cubic lattice points, alternately as in Fig. 5.21a, with each Na surrounded by 6 neighboring Cl, and *vice versa*. Similarly, the nuclei of CsCl (and similar compounds CsBr, TlCl,…) sit at the lattice sites of a bcc lattice, alternately as in Fig. 5.21b, with each Cs at the center of

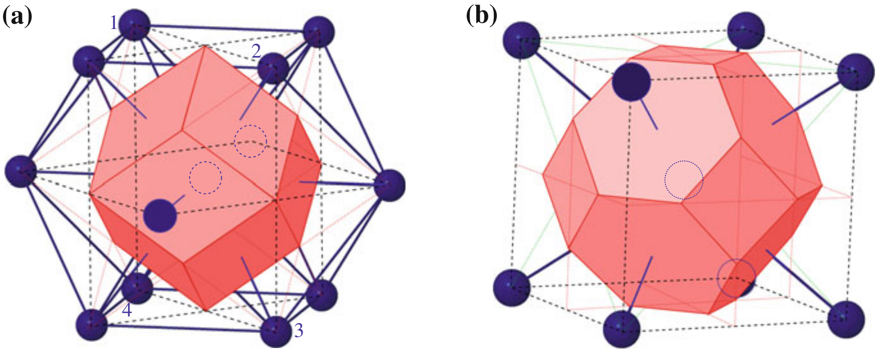


Fig. 5.20 **a** The Wigner-Seitz cell of the fcc lattice (a *rhombic dodecahedron*). A face of the conventional cubic cell of side a is the *square* whose corners are points 1–4. The lattice site around which the Wigner-Seitz cell is constructed (*dashed circle*) sits at the center of this square. **b** The Wigner-Seitz cell of the bcc lattice (a *truncated octahedron*). Each regular hexagon bisects a segment joining the conventional cell center to each vertex of the cube. Each truncation square bisects a segment joining the cell center to the center of one of the 6 neighboring cubes (*not drawn*). The corner of each truncation square sits midway between the center and the edge of a conventional-cell cube face

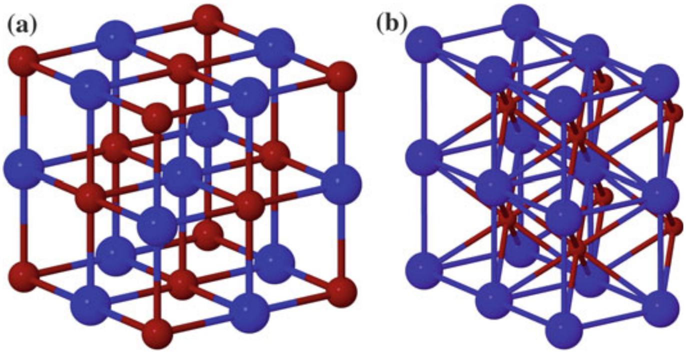


Fig. 5.21 **a** The sodium-chloride structure. **b** The cesium-chloride structure. *Large* and *small balls* represent anions and cations. In the NaCl crystal, the ions of each kind form interpenetrating fcc lattices. In the CsCl crystal, the ions of each kind form interpenetrating simple-cubic lattices

a cube with 8 neighbor Cl at its corners, and *vice versa*. To describe such structures, it suffices to recognize the true periodicity of the lattice. For example the NaCl structure can be viewed as a fcc Bravais lattice containing two atoms per cell: a Na atom at $\mathbf{0}$ and a Cl atom at the center of the fcc primitive cell $(\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3)/2 = (\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}})\, a/2$. This leads us to the necessity of introducing a *basis*, i.e. a list of atoms with their positions within a primitive cell. A *Bravais lattice plus a basis* define completely a *crystal structure*.

A basis is sometimes needed even when all atoms in the crystal are chemically equal. For example, Fig. 5.22 illustrates the 2D honeycomb net (all sites hosting "equal" atoms), which is not a simple Bravais lattice, because two neighboring atoms

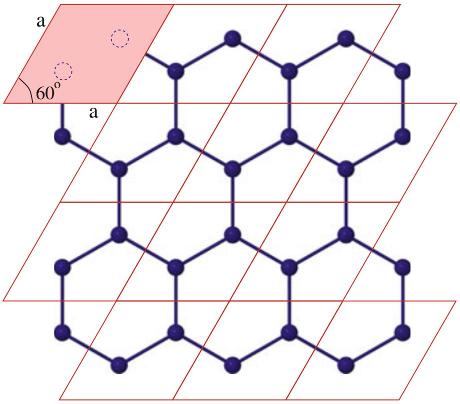


Fig. 5.22 The honeycomb crystal structure is a 2D triangular net to which one third of the points has been removed. It is not a simple Bravais lattice, since the geometric environment of two neighboring atoms is inequivalent. The honeycomb net can be described as a Bravais lattice with two primitive vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ of equal length a, separated by a 60° angle, with a basis composed of, e.g., $\mathbf{0}$ and $(\mathbf{a}_1 + \mathbf{a}_2)/3$. Graphene is a 2D honeycomb crystal of carbon, with $a \simeq 244 - 246\,\text{pm}$

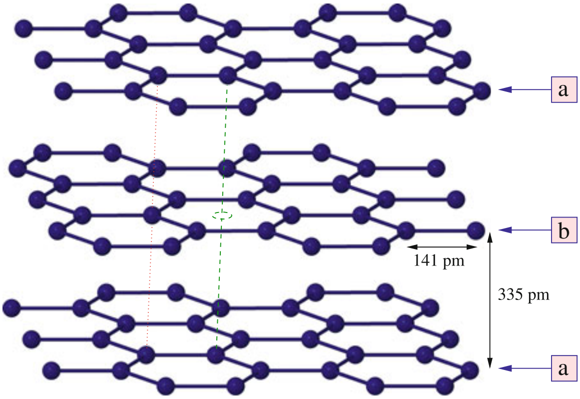


Fig. 5.23 The structure of the graphite form of carbon: an alternating stack of 2D honeycomb nets (graphene planes, Fig. 5.22). This crystal structure is a 3D hexagonal lattice with a basis of 4 atoms per primitive cell: 2 atoms in plane "a" plus 2 in plane "b"

are *geometrically inequivalent*. The honeycomb net is the 2D crystal structure of *graphene*, one atomically thin layer of the *graphite* form of carbon. The structure of graphite, shown in Fig. 5.23, is a "vertical" alternating stack of graphene sheets. The graphite structure is then a simple hexagonal lattice with a basis of 4 atoms per primitive cell. The hcp structure is also an hexagonal Bravais lattice, but with a 2-atom basis, see Fig. 5.24. The equilibrium structure of many elements, e.g. beryllium, titanium, zinc, zirconium, is hcp, although usually deviating from ideal close-packed structure, which should have $c = a\sqrt{8/3}$. Similarly, the diamond structure (of C-diamond, Si, Ge, and α-Sn) drawn in Fig. 5.25 is a fcc lattice with a 2-atom basis.

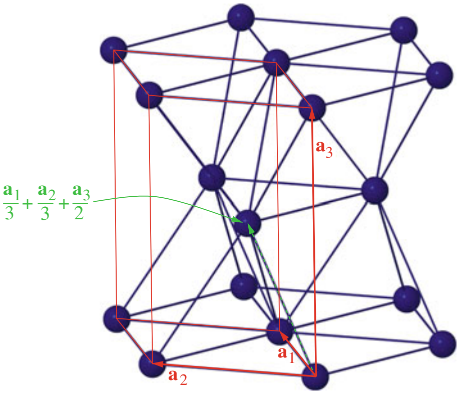


Fig. 5.24 The hcp structure, with a *highlighted parallelepiped* primitive cell. This crystal structure involves a basis consisting of 2 atoms, e.g. one at $\mathbf{0}$ and one at $\mathbf{a}_1/3+\mathbf{a}_2/3+\mathbf{a}_3/2$. The hexagonal-lattice primitive vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ of equal length a form a 60° angle; $\mathbf{a}_3$, perpendicular to the $(\mathbf{a}_1, \mathbf{a}_2)$ plane, has length c, which for the ideal close-packed structure is $c = a\sqrt{8/3}$. This structure consists of an alternate stacking of triangular lattices, see Fig. 5.3

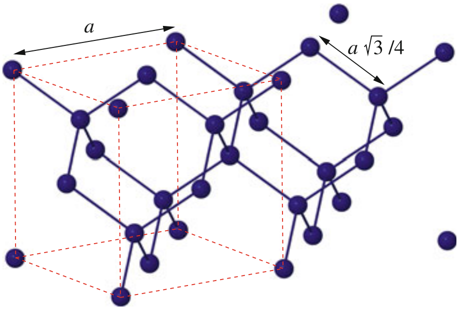


Fig. 5.25 The diamond structure consists of two interpenetrating fcc Bravais lattices, displaced along the body diagonal of the cubic cell by one quarter of the length of this diagonal. It can be regarded as a fcc lattice with a basis of two-atoms at $\mathbf{0}$ and $(\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3)/4 = (\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}})a/4$

Each atom sits at the center of a tetrahedron, surrounded by four nearest neighbors (see also Fig. 5.4).

Given the possibility that several atoms belong to each cell, one often finds it more convenient to adopt a *conventional cell* containing several *equal* atoms, rather than the primitive unit cell of the Bravais lattice. For example, fcc and bcc are often conveniently described in terms of the underlying nonprimitive simple-cubic cell, of larger volume than the primitive cells of Fig. 5.18. The fcc lattice is then viewed as a simple-cubic Bravais lattice with a 4-atom basis $[\mathbf{0}, (\hat{\mathbf{y}} + \hat{\mathbf{z}})a/2, (\hat{\mathbf{z}} + \hat{\mathbf{x}})a/2$, and $(\hat{\mathbf{x}} + \hat{\mathbf{y}})a/2]$, while the bcc is viewed as a simple-cubic lattice with a 2-points basis $[\mathbf{0}$ and $(\hat{\mathbf{x}}+\hat{\mathbf{y}}+\hat{\mathbf{z}})a/2]$. All results must coincide when this conventional "lattice with basis" formalism is used in place of the genuine description in terms of a Bravais lattice.

As a convenient picture, we introduced the Bravais lattice as an infinite array of discrete geometric points. In fact, the precise mathematical meaning of a Bravais lattice is a *group of translations*, which transform the corresponding array of points back into itself.[7] We leave the details of the classification of space and point groups of 3D crystals[8] to specific solid-state courses, and only note that extra symmetries often induce extra degeneracies of the electronic and vibrational states of the crystal.

5.1.2 The Reciprocal Lattice

The *Fourier transform of a periodic function includes only discrete "frequencies"*: this statement is the key to the introduction of a reciprocal lattice. We first illustrate this concept in 1D. By definition, a periodic function $f(x)$ of period a (the lattice spacing) satisfies $f(x) = f(x - na) = f(x - R)$ [$R = na$ is a lattice vector, as in Eq. (5.2)]. Then, in its Fourier expansion

$$f(x) = F^{-1}[\tilde{f}](x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \tilde{f}(k) dk, \tag{5.3}$$

all Fourier components $\tilde{f}(k)$ vanish except those whose $\exp(ikx)$ has the same periodicity as $f(x)$. The corresponding k must satisfy $\exp(ikx) = \exp[ik(x - a)]$, i.e.

---

[7] Any vector $\mathbf{R}$ can be also viewed as a translation operator $T_\mathbf{R}$ such that $T_\mathbf{R}\mathbf{r} = \mathbf{R} + \mathbf{r}$, for any point $\mathbf{r}$. It is easily verified that the set of all lattice translations $\{T_\mathbf{R}\}$ of a Bravais lattice is closed for composition ($T_\mathbf{R} T_{\mathbf{R}'} = T_{\mathbf{R}+\mathbf{R}'}$), it contains a neutral element ($T_\mathbf{0}$), and for each $T_\mathbf{R}$ there exists an inverse element ($T_{-\mathbf{R}}$) such that the composition of $T_\mathbf{R}$ and its inverse yields the neutral element. This means that all lattice translations $\{T_\mathbf{R}\}$ form a *group of geometric transformations*. As $\mathbf{R} + \mathbf{R}' = \mathbf{R}' + \mathbf{R}$, this group is *Abelian*, i.e. commutative. When the Hamiltonian has the entire symmetry of this group of discrete translations, its eigenstates are simultaneous eigenstates of all group operations, i.e. they are labeled by the group *irreducible representations*, see Sect. B.4. The purpose of the following section is precisely to identify these irreducible representations: we shall see that their structure is very general and not especially complicated.

[8] In addition to the discrete translations, most lattices and crystal structures have *extra symmetries*. For example, a simple-cubic lattice transforms back into itself if rotated around a lattice point by 90° around the $\hat{\mathbf{x}}$ axis, or by 120° around a body diagonal direction such as $\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}$. These extra transformations extend the group of discrete translations $\{T_\mathbf{R}\}$ outlined above. The full symmetry group of the crystal (*space group*) is a proper combination of the *point group* (a finite group of rigid rotations and reflections about one point) and the lattice group of discrete translations $\{T_\mathbf{R}\}$. Back in the first half of the 19th century it was recognized that only 7 *inequivalent point groups* could occur for 3D crystals. Two groups are equivalent if they contain the same type of symmetry operations (e.g., rotations by a certain angle, discrete translations up to suitable scaling factors). In particular, there do not exist point groups including fivefold axes (rotations by $2\pi/5 = 72°$), or sevenfold or higher-order axes, as they could not replicate infinitely in space.

These 7 point groups combine differently with the translations to form 14 *inequivalent Bravais lattices*. The introduction of a basis into the Bravais lattices can reduce the global symmetry of the replicated objects in the primitive cell, and thus the overall space group. When combined with all possible types of basis, the different point groups become 32, rather than 7, and the space groups 230, rather than 14.

$\exp(-ika) = 1$, i.e. $ka = 2\pi l$, i.e. $k = l \cdot 2\pi/a$, for any integer $l = 0, \pm 1, \pm 2, \ldots$ We indicate those special k-values compatible with the lattice periodicity with the notation

$$G = G(l) = l \cdot \frac{2\pi}{a}. \tag{5.4}$$

As at any value of $k \neq G$ which does not respect the lattice periodicity the Fourier component $\tilde{f}(k)$ vanishes, the Fourier expansion (5.3) can be written as a discrete Fourier series

$$f(x) = \sum_G e^{iGx} \tilde{f}(G), \tag{5.5}$$

with coefficients

$$\tilde{f}(G) = \frac{1}{a} \int_0^a \exp(-iGx) f(x) dx. \tag{5.6}$$

When periodic functions of period a are represented in Fourier space, the G points acquire therefore a special role among all k's. According to Eq. (5.4), the G points form a regular lattice, of unit vector $2\pi/a$, in k space. Apart from the physical dimensions of inverse length rather than length, the k space is similar to the x space, thus the lattice of G points holds all the properties of a Bravais lattice on its own: the lattice of G points of Eq. (5.4) is called *reciprocal lattice*. By definition, *direct-lattice* points R and reciprocal-lattice points G satisfy

$$e^{iRG} = e^{inal2\pi/a} = e^{i2\pi nl} = 1. \tag{5.7}$$

The simple 1D example introduced here can be generalized to the 3D case relevant for real-life crystals. A function $f(\mathbf{r})$ has the periodicity of a Bravais lattice if $f(\mathbf{r}) = f(\mathbf{r}-\mathbf{R})$ for any lattice vector $\mathbf{R} = n_1\mathbf{a}_1+n_2\mathbf{a}_2+n_3\mathbf{a}_3$, with integer n_j as in Eq. (5.2). Then, the only nonzero components in its Fourier expansion satisfy $\exp(i\mathbf{G} \cdot \mathbf{r}) = \exp[i\mathbf{G} \cdot (\mathbf{r} - \mathbf{R})]$ for all $\mathbf{R}$ in the direct lattice. This relation is satisfied for all $\mathbf{G}$ such that

$$e^{i\mathbf{R}\cdot\mathbf{G}} = 1. \tag{5.8}$$

In words, *the vectors* $\mathbf{G}$ *of the reciprocal lattice are all the* $\mathbf{k}$ *vectors whose associated plane wave has the periodicity of the direct Bravais lattice*. In particular, taking the primitive vectors $\mathbf{a}_j$ for $\mathbf{R}$, one finds that the $\mathbf{G}$ vectors are all the vectors of the type

$$\mathbf{G} = \mathbf{G}(l_1, l_2, l_3) = l_1\mathbf{b}_1 + l_2\mathbf{b}_2 + l_3\mathbf{b}_3 \tag{5.9}$$

(l_j are integers), with

$$\mathbf{b}_1 = 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{V_\mathrm{c}}, \quad \mathbf{b}_2 = 2\pi \frac{\mathbf{a}_3 \times \mathbf{a}_1}{V_\mathrm{c}}, \quad \mathbf{b}_3 = 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{V_\mathrm{c}}, \tag{5.10}$$

where $V_\mathrm{c} = \mathbf{a}_1 \times \mathbf{a}_2 \cdot \mathbf{a}_3$ is the volume of the unit cell. As above, the $\mathbf{G}$ vectors form a Bravais lattice, of which Eq. (5.10) yields a set of primitive unit vectors. Note that as $\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}$, the dot product in Eq. (5.8) $\mathbf{R} \cdot \mathbf{G} = 2\pi(n_1 l_1 + n_2 l_2 + n_3 l_3)$. Any other vector $\mathbf{k} \neq \mathbf{G}$ is associated to a plane wave which does not respect the lattice periodicity, thus the corresponding Fourier component $\tilde{f}(\mathbf{k})$ vanishes. Like in 1D, the Fourier expansion of the periodic function is a discrete Fourier summation over the reciprocal lattice

$$f(\mathbf{r}) = \sum_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}} \tilde{f}(\mathbf{G}), \tag{5.11}$$

with coefficients

$$\tilde{f}(G) = \frac{1}{V_\mathrm{c}} \int_{V_\mathrm{c}} \exp(-i\mathbf{G} \cdot \mathbf{r})\, f(\mathbf{r}) d^3r, \tag{5.12}$$

where the integration is carried out over a single unit-cell volume. The volume $\mathbf{b}_1 \times \mathbf{b}_2 \times \mathbf{b}_3$ of the reciprocal-lattice primitive cell is $(2\pi)^3/V_\mathrm{c}$. The Wigner-Seitz cell of the reciprocal lattice is called *first Brillouin zone* (BZ).

By applying the transformations (5.10), it is easy to verify that the reciprocal lattice of a simple-cubic lattice of side a is another simple-cubic lattice of side $2\pi/a$. The fcc lattice of conventional cell of side a has as reciprocal lattice a bcc of conventional side $4\pi/a$. Conversely, the bcc lattice of conventional cell of side a has a fcc reciprocal lattice[9] of conventional cell side $4\pi/a$. As a consequence, the first BZ of the fcc lattice has the shape of the bcc Wigner-Seitz cell (Fig. 5.20b), and that of the bcc lattice has the shape of the fcc Wigner-Seitz cell (Fig. 5.20a). It is a useful exercise to determine the reciprocal lattice of the hexagonal lattice starting from the conventional primitive vectors of Fig. 5.24.

$\mathbf{G}$ vectors in the reciprocal lattice select the plane waves $e^{i\mathbf{G}\cdot\mathbf{r}}$ with the direct-lattice periodicity. Consider the constant-wave surfaces of a given plane wave, namely its "wave fronts", for example those fixed by $e^{i\mathbf{G}\cdot\mathbf{r}} = 1$: these surfaces are a family of parallel planes, perpendicular to $\mathbf{G}$ and separated by one wavelength $\lambda = 2\pi/|\mathbf{G}|$. Some of these planes pass through the lattice points. In particular all these planes pass through lattice points if the integer indexes l_1, l_2, and l_3 have no common nontrivial divisors. Otherwise, for $\mathbf{G}$ given by nl_1, nl_2, nl_3, one plane out of n goes through lattice points. The integer indexes l_1, l_2, and l_3 are called *Miller indexes* of the family of lattice planes. These indexes are inversely proportional to the intercepts of these planes with the crystal primitive directions. The standard notation to indicate planes and directions in $\mathbf{k}$ space is $(l_1\, l_2\, l_3)$—no commas, with minuses conventionally represented by overbars, as in $(2\ {-1}\ 0) \to (2\ \bar{1}\ 0)$. Figure 5.26 illustrates the relation of a few popular plane families with the real-space cubic cell. Traditionally, for all cubic lattices, including fcc and bcc, lattice planes are labeled relative to the

---

[9] This is a consequence of the general fact that the reciprocal of the reciprocal lattice is the original lattice. This can be verified by applying twice the transformations (5.10), or even more simply by observing that the roles of $\mathbf{R}$ and $\mathbf{G}$ in Eq. (5.8) can be exchanged.

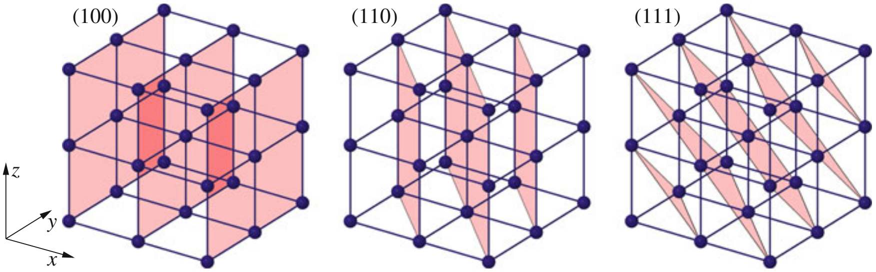


Fig. 5.26 A $\mathbf{G}$ vector identifies a well-defined family of parallel lattice planes, of constant phase $\mathbf{G} \cdot \mathbf{r}$. of the plane wave $e^{i\mathbf{G}\cdot\mathbf{r}}$. These planes are perpendicular to $\mathbf{G}$ and separated by a distance $\frac{2\pi}{|\mathbf{G}|}$. The labels report the standard notation for the Miller indexes $(l_1\, l_2\, l_3)$ identifying families of lattice planes in cubic lattices: (1 0 0), (1 1 0), and (1 1 1)

conventional cubic directions $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$, not relative to the primitive unit vectors of Fig. 5.16.[10]

5.1.3 Diffraction Experiments

Diffraction of "wave probes" is the main quantitative source of structural data about solids. Electrons of energy 1.5–150 eV, electromagnetic radiation of 1–10 keV (see Fig. 1.2), and neutrons of 1–100 meV: the wavelength of all these "waves" fits the typical unit-cell size of not too complicated crystals (of the order of a few typical interatomic distances, 0.1–1 nm). The periodic density modulation in solids can and does diffract waves of these three kinds. Of the 3 considered wave probes, electrons interact very strongly with matter, and are thus sensitive to few topmost surface layers only. If the energy of X-rays is chosen off-resonance from all core excitations of the atoms in the material (see Fig. 2.22), then their penetration depth in solids easily exceeds thousands of unit-cell lengths, sufficient to generate sharp bulk diffraction patterns. Slow neutrons are even better fit for structural diffraction studies, as they only interact mostly with the nuclei, and boast penetration depths of several centimeters. A sufficiently small sample guarantees that the total probability of probe-sample interaction is small, so that most probing radiation goes unscattered

---

[10] The reciprocal lattice helps us in identifying all different irreducible representations (see Sect. B.4) of the group of the discrete direct-lattice translations. These representations, labeled by an arbitrary $\mathbf{k}$ vector, are 1-dimensional and the corresponding character of a group operation $T_{\mathbf{R}}$ is simply $\exp(-i\mathbf{k} \cdot \mathbf{R})$. Two irreducible representation labeled by $\mathbf{k}$ and $\mathbf{k}'$ have all equal characters (thus are in fact the same representation) whenever $\mathbf{k} - \mathbf{k}' = \mathbf{G}$ for some $\mathbf{G}$ in the reciprocal lattice. Indeed, $\exp(-i\mathbf{k} \cdot \mathbf{R}) = \exp(-i[\mathbf{k}' + \mathbf{G}] \cdot \mathbf{R}) = \exp(-i\mathbf{k}' \cdot \mathbf{R}) \exp(-i\mathbf{G} \cdot \mathbf{R}) = \exp(-i\mathbf{k}' \cdot \mathbf{R})$, using Eq. (5.8). Accordingly, all different irreducible representations of the discrete translational group are labeled by all $\mathbf{k}$ points within one primitive zone of the reciprocal lattice, e.g. the first BZ.

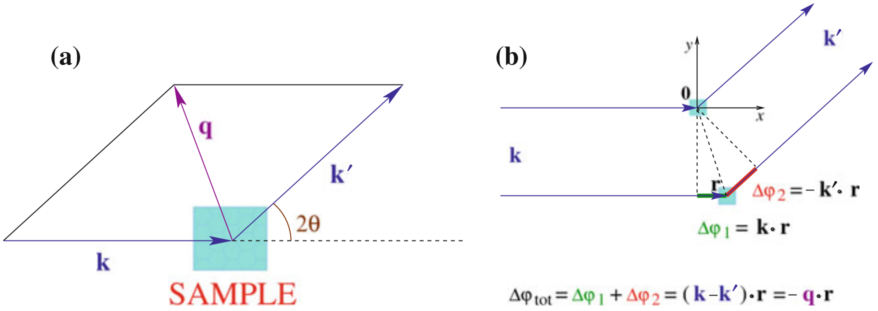


Fig. 5.27 **a** In an elastic-scattering experiment ($|\mathbf{k}'| = |\mathbf{k}|$), the transferred wave vector $\mathbf{q} = \mathbf{k}' - \mathbf{k}$ holds information about the scattering angle 2θ and radiation wavelength $\lambda = 2\pi/|\mathbf{k}|$. **b** The "far-field" wave scattered by a point-like element of matter at $\mathbf{r}$ is dephased by $-\mathbf{q} \cdot \mathbf{r}$ compared to the one at the origin $\mathbf{0}$, due to the extra path length $\hat{\mathbf{q}} \cdot \mathbf{r}$ emphasized by bold segments. The evaluation of the total interfering wave from a continuous distribution leads to a Fourier-transform "summation"

through the sample, a small fraction scatters once, and almost none scatters twice or more. Under these conditions, understanding scattering is straightforward.

For simplicity, the incoming beam is assumed to be produced by a monochromatic source placed at a large distance from the sample, so that it is characterized by a well-defined wave vector $\mathbf{k}$ (or momentum $\hbar\mathbf{k}$). The detector is also very remote from the sample, so that it detects outgoing radiation scattered elastically to another well-defined wave vector $\mathbf{k}'$ (Fig. 5.27a). As analyzed quantitatively in the theory of scattering, every infinitesimal volume $d^3\mathbf{r}$ of a continuous distribution of matter scatters radiation in proportion to the number of scatterers $n(\mathbf{r})d^3\mathbf{r}$ locally present. As suggested around Eq. (2.45) and sketched in Fig. 5.27b, the total rate of transition $\gamma_{\mathbf{k}\,\mathbf{k}'}$ is proportional to the square modulus of the matrix element $\int e^{-i\mathbf{k}'\cdot\mathbf{r}} n(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} d^3\mathbf{r} = \int e^{-i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}} n(\mathbf{r}) d^3\mathbf{r} \propto F[n](\mathbf{k}' - \mathbf{k}) = \tilde{n}(\mathbf{k}' - \mathbf{k})$, namely the Fourier transform of the density of scatterers. These observations suggest that the elastic scattering rate of the incoming $\mathbf{k}$ wave radiation into the $\mathbf{k}'$ direction is *proportional to the square modulus of the Fourier transform* $\tilde{n}$ of the appropriate *density* n, evaluated at the *transferred wave vector* $\mathbf{q} = \mathbf{k}' - \mathbf{k}$ (Fig. 5.27a). For neutrons the relevant density is the nuclear-matter density $n_{\text{nuc}}(\mathbf{r})$ while for X-rays it is the electronic density $n_{\text{el}}(\mathbf{r})$:

$$I_{\substack{\text{neutr}\\ \text{X-rays}}}(\mathbf{q}) \propto \left| F[n_{\substack{\text{nuc}\\ \text{el}}}](\mathbf{q}) \right|^2 \equiv \left| \tilde{n}_{\substack{\text{nuc}\\ \text{el}}}(\mathbf{q}) \right|^2 . \tag{5.13}$$

As a nucleus is essentially point-like, if viewed at atomic resolution, the Fourier transform of the density distribution of one nucleus is basically flat: thus each atom scatters neutrons essentially independently of $\mathbf{q}$, see Fig. 5.28a. When *two* atoms diffuse neutrons, the scattered matter wave *interfere*: interference leads to intensity reinforcement and reduction in alternating directions. Constructive interference occurs whenever the two source-atom-detector path lengths (Fig. 5.27b) differ by an

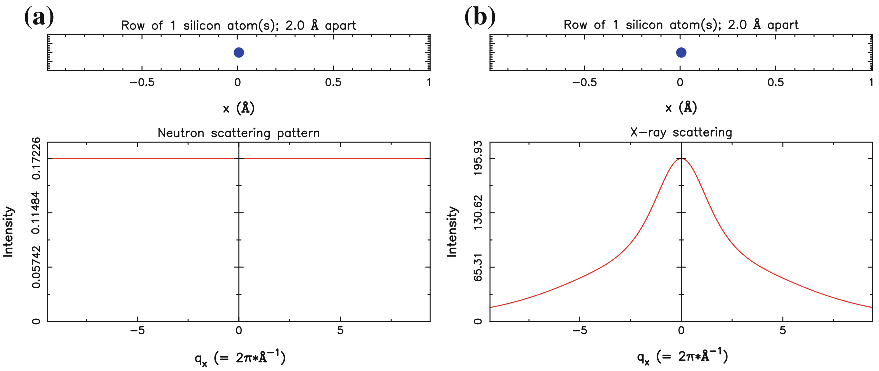


Fig. 5.28 Comparison of **a** neutron and **b** X-ray elastic scattering by a single Si atom. q_x represents the x-component of the change in wave vector $\mathbf{k}' - \mathbf{k}$ of the scattered radiation. Intensity is proportional to the square modulus of the Fourier transform of the appropriate density. Neutrons probe the nuclear-matter density, which is concentrated in lumps of typical size ∼1 fm, practically similar to a Dirac-delta distribution $\delta(\mathbf{r})$: its Fourier transform is a constant, essentially independent of $\mathbf{q}$ (until a huge $|q| \gtrsim 10^4$ Å^{-1}, irrelevant for diffraction experiments of crystals). X-rays probe the Fourier transform of the total atomic electronic charge density, which lumps in a region size of ∼1 Å: thus the atomic form factor varies significantly over a q_x range of few Å^{-1} (Generated using the software of Ref. [36])

integer multiple of the radiation wavelength λ. Quantitatively, the square modulus of the Fourier transform $\tilde{n}_{\text{nuc}}(\mathbf{q})$ of two equal point-like objects at $\mathbf{R}_1$ and $\mathbf{R}_2$

$$\begin{aligned} |\tilde{n}_{\text{nuc}}(\mathbf{q})|^2 &\propto |\exp(-i\mathbf{R}_1 \cdot \mathbf{q}) + \exp(-i\mathbf{R}_2 \cdot \mathbf{q})|^2 \\ &= \left| \exp\left(-i\, \frac{\mathbf{R}_1 + \mathbf{R}_2}{2} \cdot \mathbf{q}\right) 2\cos\left(\frac{\mathbf{R}_1 - \mathbf{R}_2}{2} \cdot \mathbf{q}\right)\right|^2 \\ &= 2\left[1 + \cos\left((\mathbf{R}_1 - \mathbf{R}_2) \cdot \mathbf{q}\right)\right] = 2\left[1 + \cos\left(a\, q_x\right)\right] \end{aligned} \qquad (5.14)$$

shows characteristic oscillations (Fig. 5.29a, c). Assume that $\mathbf{R}_1 - \mathbf{R}_2 = a\hat{\mathbf{x}}$: constructive interference yields maximum intensity in the $\mathbf{q}$ directions characterized by a projection q_x along the line joining the two nuclei such that $(\mathbf{R}_1 - \mathbf{R}_2) \cdot \mathbf{q} \equiv a q_x = 2\pi \times$ an integer l_1. The role of interatomic separation a on the interference pattern of the scattered waves is illustrated by the comparison of Fig. 5.29a, c: an increased separation in real space produces closer interference maxima in $\mathbf{q}$ space. Scattered intensity is independent of the q_y, q_z components in the plane (not shown in Fig. 5.29) orthogonal to the line joining the two atoms (see also Fig. 5.32a for a 2D example).

Neutron scattering from a crystal arises from the interference of the waves scattered coherently by a regular array of point scatterers. The Fourier transform in Eq. (5.13) is again a discrete sum $\sum_{n_1} e^{-i q_x n_1 a}$, recalling the Fraunhofer theory of light diffracted by an optical grating with narrow slits. Figure 5.30 shows the intensity scattered by short regular chains of atoms: despite the smallness of such 1D "crystals", sharp intense *diffraction structures* emerge at regular $\mathbf{q}$ directions

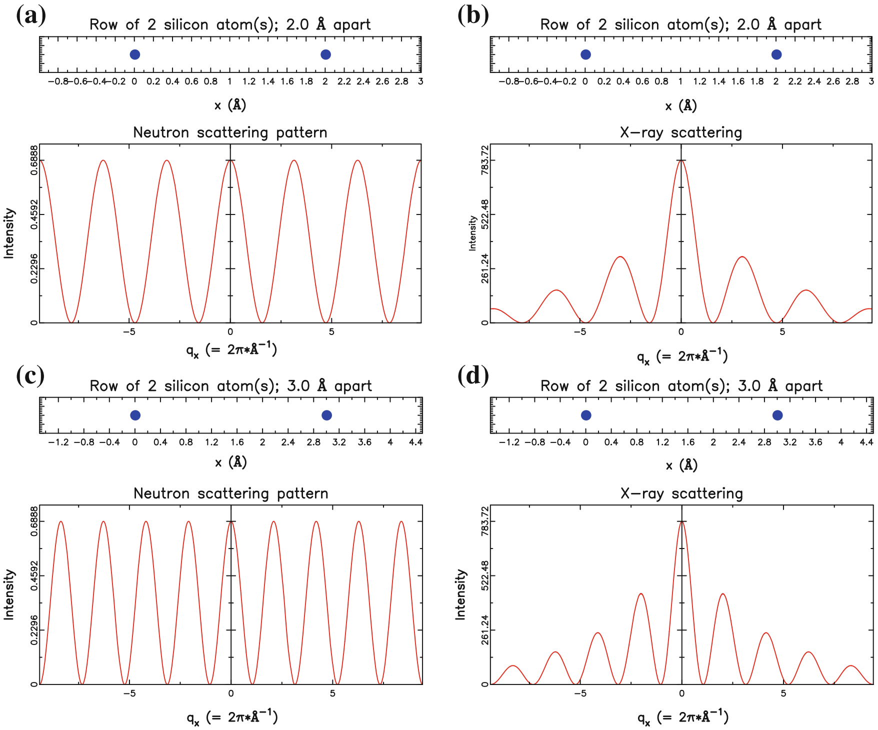


Fig. 5.29 **a**, **c** Neutron and **b**, **d** X-ray scattering by two Si atoms. The patterns shows the characteristic interference periodicity $\frac{2\pi}{a}$, where $a = |\mathbf{R}_1 - \mathbf{R}_2|$ is the distance between the two nuclei, (**a**, **b**) $a = 2\,\text{Å}$ and (**c**, **d**) $a = 3\,\text{Å}$. According to Eq. (5.18), the X-rays pattern is the product of the neutron pattern (uniquely determined by the atomic positions) times the atomic form factor, carrying information on the charge distribution of the single atom. Note that, due to the enlarged interatomic separation a, the interference fringes in (**c**, **d**) are closer than in (**a**, **b**), while the enveloping atomic form factor remains unchanged (Generated using the software of Ref. [36])

characterized by $e^{-iq_x n_1 a} = 1$ for all n_1. In between these strong Bragg diffraction peaks, the intensity of the weak secondary interference peaks decreases rapidly as the number of atoms increases. Already for 30 atoms these weak structures become almost invisible (Fig. 5.30c, d), and disappear in the limit of a macroscopically large regular crystal.[11] The values of $\mathbf{q}_x$ for the Bragg maxima of diffracted intensity, with all phase factors $q_x\, n_1\, a =$ an integer multiple of 2π are precisely the 1D reciprocal lattice vectors $q_x = l_1 \times 2\pi/a$. The same occurs also for 2D and 3D periodicity: we conclude that Bragg *diffraction peaks occur for transferred wave vector equal to reciprocal-lattice points* $\mathbf{q} = \mathbf{G}$.

---

[11] The modest regularity sufficient to produce diffracted peaks grants the possibility to generate detectable diffraction patterns even in the event that the coherence length of the probing radiation available in the lab extends to several lattice cells spacings only, much shorter than the whole sample size.

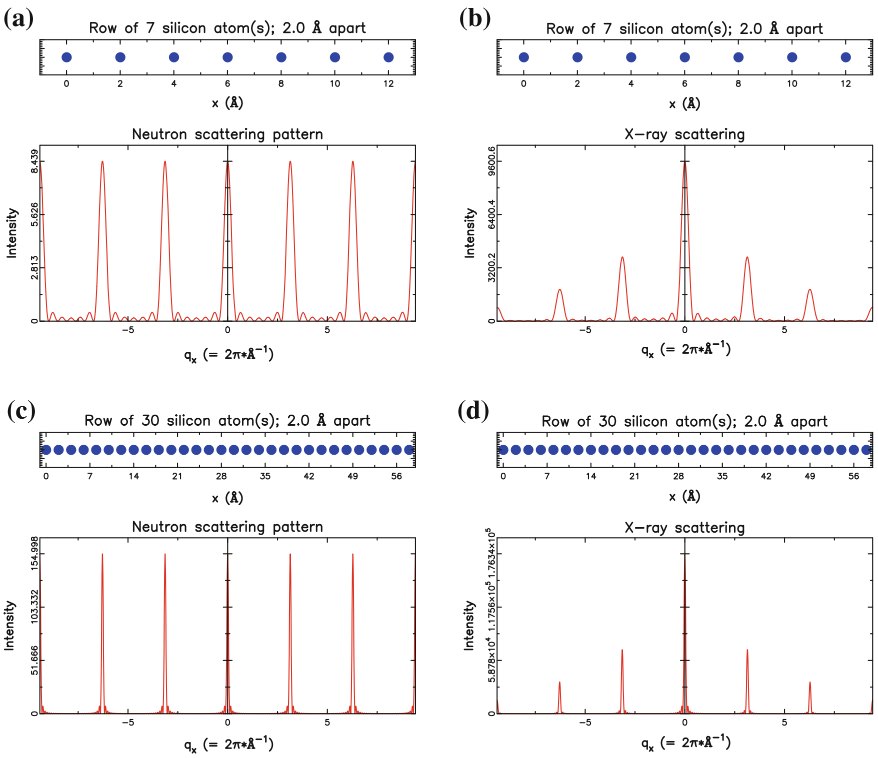


Fig. 5.30 Neutron and X-rays diffraction patterns but generated by 7 (**a**, **b**) and 30 (**c**, **d**) atoms rather than 2 as in Fig. 5.29. The sharp diffraction peaks at distance $\frac{2\pi}{2.0}$ Å^{-1} reflect the separation $a = 2.0$ Å. The relative intensity of the subsidiary interference peaks decreases quickly as the number of atoms is increased, and vanishes in the limit of an infinitely large crystal. In the X-ray pattern, the diffraction peaks are modulated by the atomic form factor (Generated using the software of Ref. [36])

The patterns of Figs. 5.29 and 5.30 show that the X-ray diffractograms differ from neutrons diffractograms by an amplitude modulation. This is easily understood in steps:

- X-rays scattered by a single atom probe its smooth electronic density distribution $n_{\rm at}(\mathbf{r})$. According to Eq. (5.13), one atom scatters X-rays elastically with the non-trivial $\mathbf{q}$ dependence given by Fourier transform of its electronic density, called *atomic form factor* $f_{\rm at}(\mathbf{q}) \equiv \tilde{n}_{\rm at}(\mathbf{q})$, as drawn for Si in Fig. 5.28b. Figure 5.31 reports two examples of X-ray atomic form factors: these are characteristic bell-shaped functions of total weight proportional to the squared number of electrons.
- As illustrated in Sects. 3.2.1 and 3.2.2 above, chemical bonding modifies the electronic states of a molecule or a solid, so that the electronic charge density does differ from the sum of the densities of the individual atoms. However only valence electrons are involved in bonding and delocalize significantly. X-rays are scattered mostly by the (more numerous if $Z \gg 1$) core electrons, which retain an

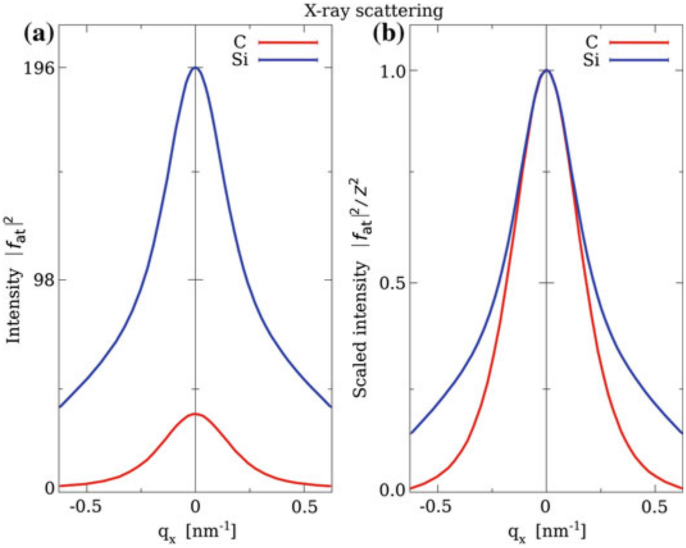


Fig. 5.31 The square moduli of the X-ray atomic form factors $|f_{at}(\mathbf{q})|^2$ of carbon ($Z = 6$) and of silicon ($Z = 14$). **a** Absolute values, showing that the total X-ray scattering increases $\propto Z^2$; **b** Intensities scaled to the $\mathbf{q} = 0$ value, showing that the form factor of the heavier atom is broader, due to the sharper charge localization of its inner-core electrons

atomic-like charge distribution. Accordingly, it is a fair approximation to assume that the electronic distribution of a collection of atoms (as in a molecule, or in a solid) equals *the sum of the individual atomic electronic distributions*. For example an elementary solid consisting of many equal atoms sitting one per cell at the points $\mathbf{R}$ of a Bravais lattice has

$$n_{el}(\mathbf{r}) = \sum_{\mathbf{R}} n_{at}(\mathbf{r} - \mathbf{R}). \tag{5.15}$$

- Substitute Eq. (5.15) in the calculation of the Fourier transform:

$$\begin{aligned}\tilde{n}_{el}(\mathbf{q}) &= \sum_{\mathbf{R}} \int n_{at}(\mathbf{r} - \mathbf{R}) e^{-i\mathbf{q}\cdot\mathbf{r}} d^3\,\mathbf{r} = \sum_{\mathbf{R}} \int n_{at}(\mathbf{r}') e^{-i\mathbf{q}\cdot(\mathbf{r}'+\mathbf{R})} d^3\,\mathbf{r}' \\ &= \sum_{\mathbf{R}} e^{-i\mathbf{q}\cdot\mathbf{R}} \int n_{at}(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} d^3\,\mathbf{r} = \tilde{n}_{\text{Bravais}}(\mathbf{q})\, f_{at}(\mathbf{q}).\end{aligned} \tag{5.16}$$

Observe that the resulting amplitude is the product of two independent functions of $\mathbf{q}$: $\tilde{n}_{\text{Bravais}}(\mathbf{q}) = \sum_{\mathbf{R}} e^{-i\mathbf{q}\cdot\mathbf{R}}$ describes the scattering of point-like objects arranged at the nodes of the Bravais lattice; and the atomic form factor accounts for the scattering from a single atom, Fig. 5.28b. Accordingly, the scattered intensity

$$I_{\text{X-rays}}(\mathbf{q}) \propto |\tilde{n}_{el}(\mathbf{q})|^2 = |\tilde{n}_{\text{Bravais}}(\mathbf{q})\, f_{at}(\mathbf{q})|^2 = |\tilde{n}_{\text{Bravais}}(\mathbf{q})|^2\, |f_{at}(\mathbf{q})|^2 . \tag{5.17}$$

- Finally, observe that $n_{\rm nuc}(\mathbf{r}) \propto n_{\rm Bravais}(\mathbf{r})$, and conclude that

$$I_{\rm X-rays}(\mathbf{q}) \propto |f_{\rm at}(\mathbf{q})|^2 \, |\tilde{n}_{\rm nuc}(\mathbf{q})|^2 \propto |f_{\rm at}(\mathbf{q})|^2 \, I_{\rm neutr}(\mathbf{q}). \tag{5.18}$$

This relation clarifies the role of the atomic form factor $f_{\rm at}(\mathbf{q})$ in X-ray diffraction: X-rays scatter in the *same* $\mathbf{q}$ *directions as neutrons* of the same wavelength, but the peak *intensities are modulated multiplicatively by the squared atomic form factor*. This observation accounts for the compared diffractograms of Figs. 5.28, 5.29 and 5.30.

The 1D patterns of Figs. 5.28, 5.29 and 5.30 generalize simply to 2D and 3D. Figure 5.32 shows the 2D pattern of neutrons scattered by 2, 3, 4, and 5 equal atoms located at the vertexes of regular polygons. The fact that 3 and 4 atoms generate a Bravais lattice as an interference pattern, while the pentagonal arrangement does not, suggests that pentagonal symmetry is incompatible with repetition in space. Figure 5.33 illustrates the typical effect of geometric deformations of the lattice on the diffraction pattern, as described mathematically by Eq. (5.10). Figures 5.34, 5.35, 5.36 and 5.37 display typical neutron diffraction patterns for varied 2D geometries. It is instructive to explore other 2D patterns my means of the handy software of Refs. [36, 37].

The 3D patterns follow similar rules: diffracted beams come out in the directions where the transferred $\mathbf{q}$ matches a $\mathbf{G}$-vector of the 3D reciprocal lattice. In practice, when shining a monochromatic neutron or X-ray beam on a single crystal, one generally obtains no diffracted beams, since, for that given $\mathbf{k}$, all possible $\mathbf{k}' = \mathbf{k} + \mathbf{G}$ have lengths $|\mathbf{k}'|$ different from $|\mathbf{k}|$, and would then correspond to *inelastic* scattering. Diffracted beams are retrieved by carefully orienting the crystal until, for some $\mathbf{G}$, the incident and scattered wave numbers match:

$$|\mathbf{k}| = |\mathbf{k}'| = |\mathbf{k} + \mathbf{G}|. \tag{5.19}$$

This geometric condition is represented by the Ewald construction of Fig. 5.38a. The scattering angle, i.e. the angle between $\mathbf{k}$ and $\mathbf{k}'$ is related to $|\mathbf{k}|$ and $\mathbf{G}$ by squaring Eq. (5.19), obtaining

$$\mathbf{k} \cdot \hat{\mathbf{G}} = -\frac{|\mathbf{G}|}{2}. \tag{5.20}$$

Figure 5.38b illustrates this geometric relation: the projection $\mathbf{k} \cdot \hat{\mathbf{G}} = -|\mathbf{k}| \sin\theta$ must equal half the length of $\mathbf{G}$. Accordingly, scattering is observed at angles 2θ related to the lengths of $\mathbf{k}$ and $\mathbf{G}$ by

$$\sin\theta = \frac{|\mathbf{G}|}{2\,|\mathbf{k}|}. \tag{5.21}$$

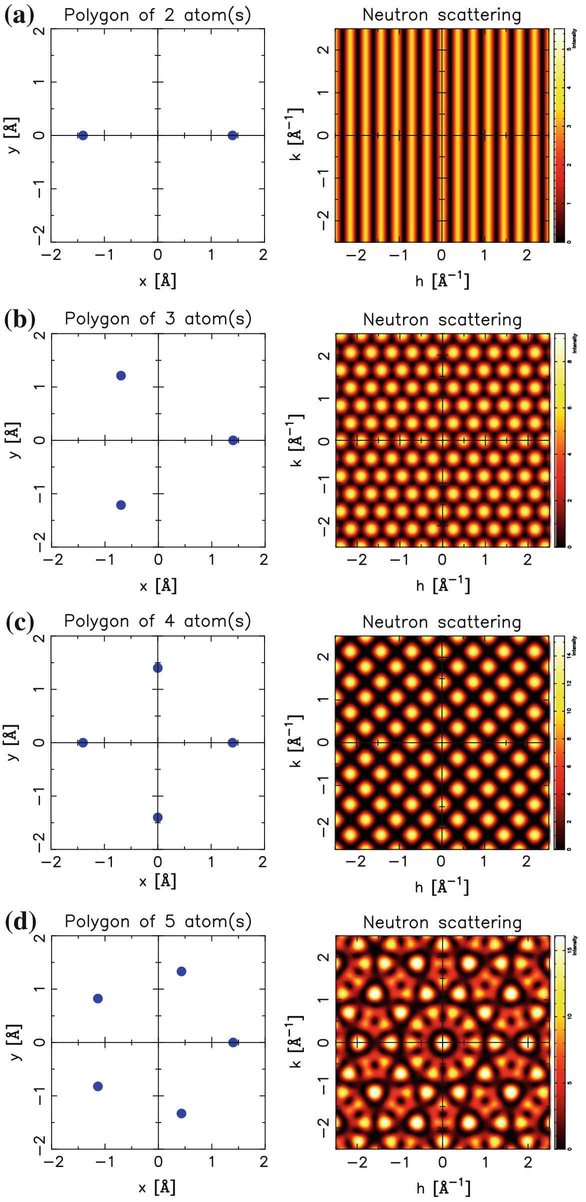


Fig. 5.32 2D neutron scattering patterns generated by regular polygons of 2–5 equal atoms. For 2 atoms (**a**), as indicated by Eq. (5.14), scattering is independent of the **q** component perpendicular to the line joining them. Note that the regular *triangle* (**b**) and *square* (**c**) generate Bravais lattices as "diffraction" patterns, while the pentagon (**d**) does not. This figure and a few others below have $h = q_1/2\pi$ and $k = q_2/2\pi$, and have been generated using the software of Ref. [36]

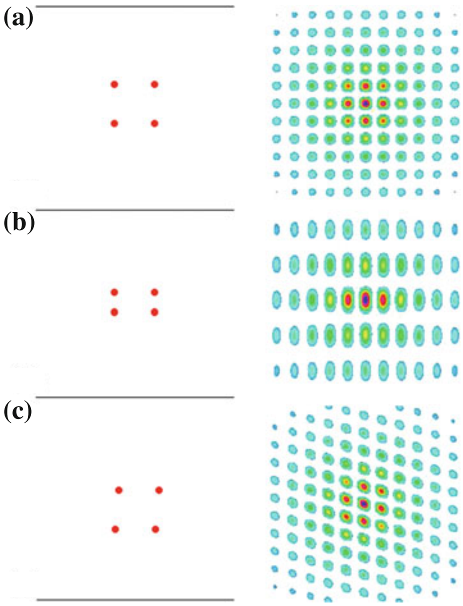


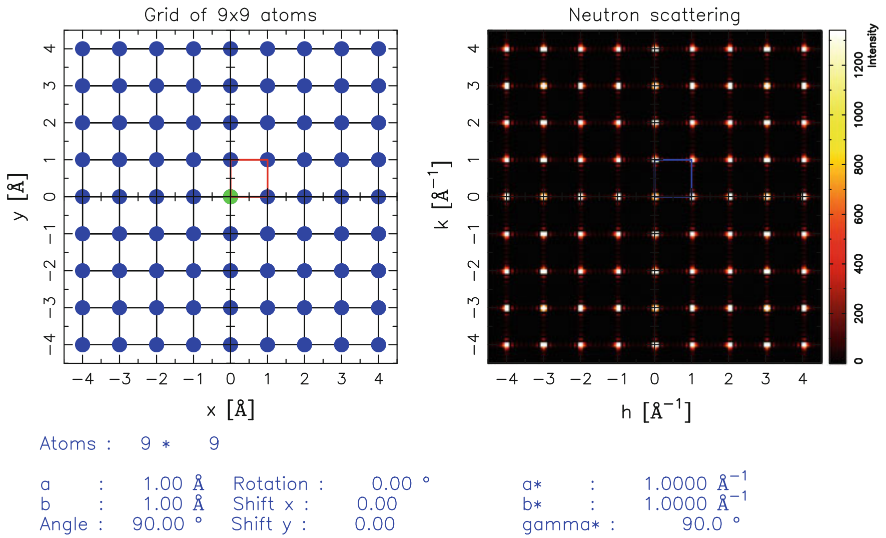
Fig. 5.33 A color-intensity plot 2D of the X-rays radiation scattered by 4 atoms.This interference pattern hints at the general features of diffraction from the *square* (**a**), *rectangular* (**b**), and oblique (**c**) nets. The diffracted spots at the reciprocal lattice points become much sharper when a large crystal, rather than 4 atoms, generates them. The inverse proportionality and orthogonality relation (5.10) of the reciprocal-lattice and direct-lattice basis is illustrated

In Sect. 5.1.2, we related **G** to a family of lattice planes (drawn in Fig. 5.38b) separated by a distance $d = n 2\pi/|\mathbf{G}|$, where n is the greatest common divisor of the Miller indexes defining **G**. On the other hand, $|\mathbf{k}|$ is connected to the radiation wavelength λ by $|\mathbf{k}| = 2\pi/\lambda$: substitution in Eq. (5.21) yields the celebrated *Bragg condition for diffraction*

$$2d \sin\theta = n\lambda. \tag{5.22}$$

According to these relations, *no diffraction occurs for* $2|\mathbf{k}| < |\mathbf{G}|$, or equivalently for $\lambda > 2d$.

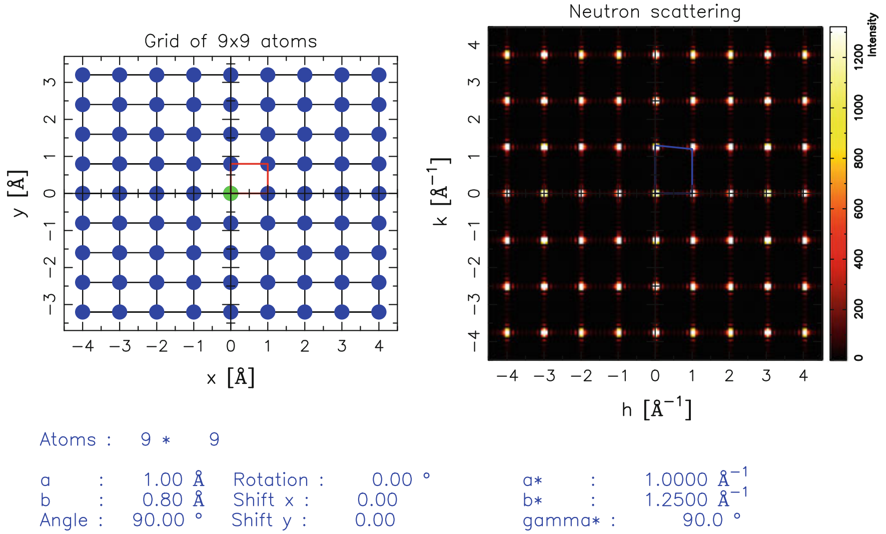
In practice, to generate diffracted beams off a single crystal, one must move the reciprocal lattice relative to the Ewald sphere until some **G** point touches the Ewald sphere, as in Fig. 5.38a. One can either vary the radiation wavelength, thus changing the Ewald sphere diameter, or else rotate the crystal sample [its reciprocal lattice rotates accordingly, see Eq. (5.10)].



```
Atoms :    9 *    9

a     :    1.00 Å   Rotation :    0.00 °        a*      :    1.0000 Å^-1
b     :    1.00 Å   Shift x  :   0.00           b*      :    1.0000 Å^-1
Angle :   90.00 °   Shift y  :   0.00           gamma*  :      90.0 °
```

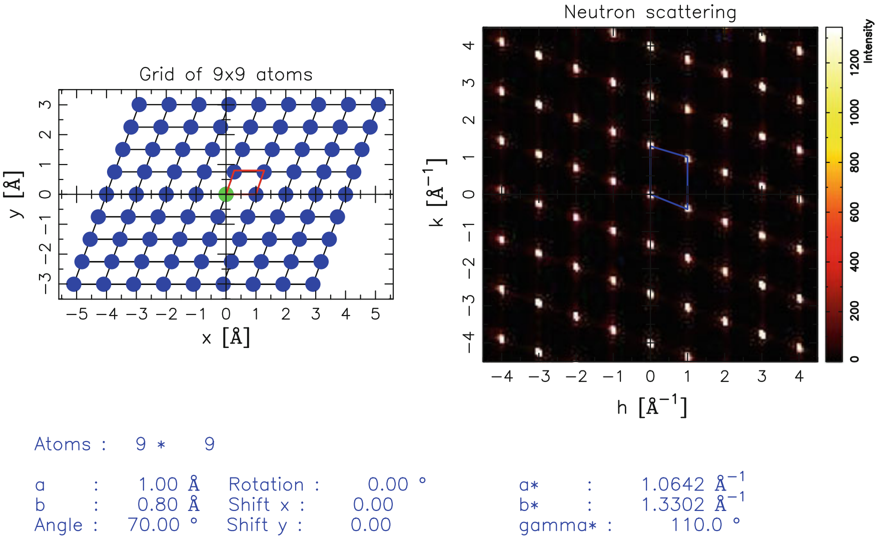
Fig. 5.34 The neutron diffraction pattern generated by a 9 × 9 *square-lattice crystal*. A *red square* marks the real-space primitive cell; a *blue square* marks the reciprocal space primitive cell. The peaks are much sharper than in the 2 × 2 example of Fig. 5.33



```
Atoms :    9 *    9

a     :    1.00 Å   Rotation :    0.00 °        a*      :    1.0000 Å^-1
b     :    0.80 Å   Shift x  :   0.00           b*      :    1.2500 Å^-1
Angle :   90.00 °   Shift y  :   0.00           gamma*  :      90.0 °
```

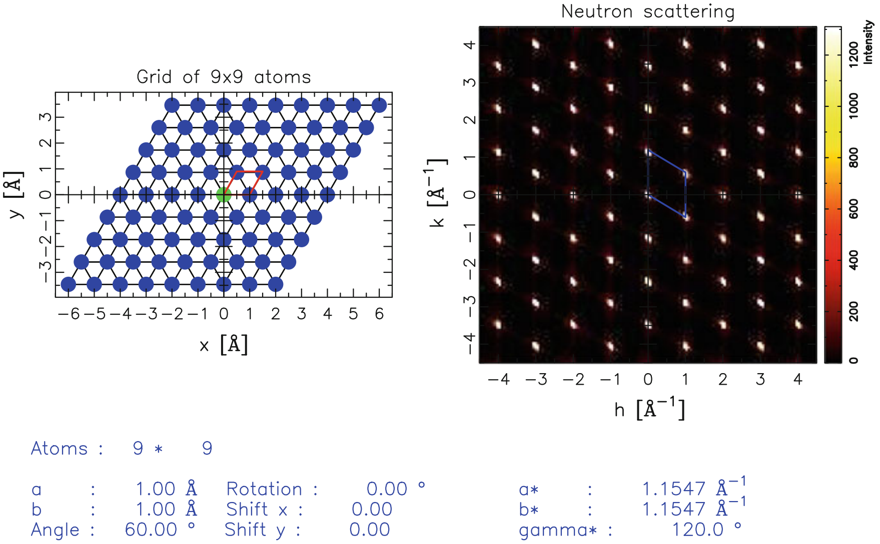
Fig. 5.35 The neutron diffraction pattern generated by a rectangular-lattice crystal



```
Atoms :    9 *     9

a      :     1.00 Å    Rotation :      0.00 °          a*      :      1.0642 Å⁻¹
b      :     0.80 Å    Shift  x :      0.00            b*      :      1.3302 Å⁻¹
Angle  :    70.00 °    Shift  y :      0.00            gamma* :       110.0 °
```

Fig. 5.36 The neutron diffraction pattern generated by an oblique-lattice crystal



```
Atoms :    9 *     9

a      :     1.00 Å    Rotation :      0.00 °          a*      :      1.1547 Å⁻¹
b      :     1.00 Å    Shift  x :      0.00            b*      :      1.1547 Å⁻¹
Angle  :    60.00 °    Shift  y :      0.00            gamma* :       120.0 °
```

Fig. 5.37 The neutron diffraction pattern generated by a triangular-lattice crystal

In the lab, it is common to characterize the structure of *powder samples*, i.e. collections of microcrystals rotated randomly in space. This uniform distribution of orientations is equivalent to averaging over all possible rotations of the reciprocal lattice. As illustrated by Fig. 5.39, the crystal scatters radiation at fixed angles 2θ [given by Eq. (5.21)] away from the incident **k**. This means that diffracted radiation

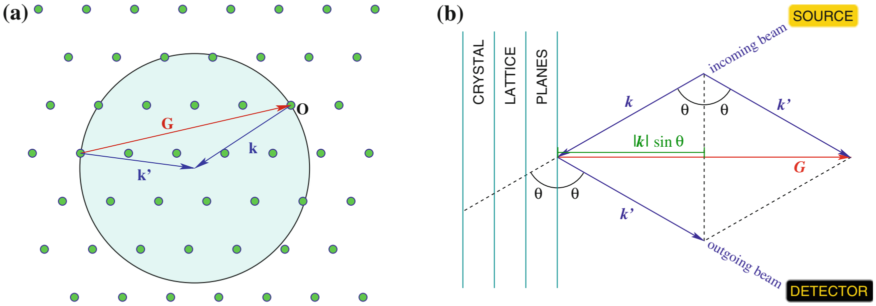


Fig. 5.38 **a** The Ewald construction. Given the incident vector $\mathbf{k}$, draw a *sphere* of radius $|\mathbf{k}|$ about the point $\mathbf{k}$ in reciprocal space (*shaded*). Diffraction peaks corresponding to reciprocal lattice vectors $\mathbf{G}$ occur only if $-\mathbf{G}$ happens to lie on the surface of the Ewald sphere, as drawn. Under this condition, radiation is diffracted to the direction $\mathbf{k}' = \mathbf{k} + \mathbf{G}$. **b** The relation between the scattering angle 2θ and the lengths of $\mathbf{k}$ and $\mathbf{G}$ vectors. The angle θ between the incident beam and the family of Bragg planes fixed by $\mathbf{G}$ yields the projection $\mathbf{k} \cdot \hat{\mathbf{G}} = -|\mathbf{k}| \sin\theta$, which is compatible with diffraction when it equals $-|\mathbf{G}|/2$

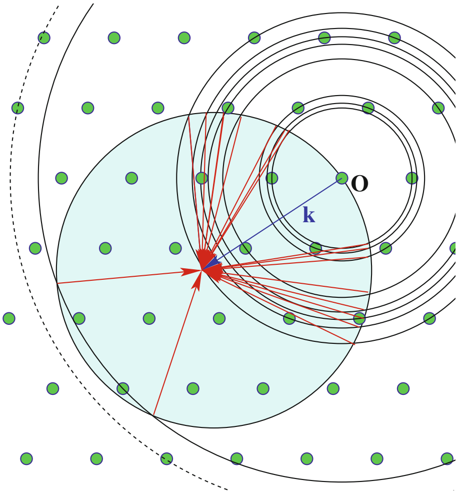


Fig. 5.39 The Ewald construction for a powder sample. The wave vector $\mathbf{k}$ of the incoming radiation defines the (*shaded*) Ewald circle in 2D (a sphere in 3D). Crystal grains take all possible rotations: the entire reciprocal lattice (*solid points*) must then be "averaged" over all angles: every $\mathbf{G}$ point covers a circle in 2D (a sphere in 3D). In 2D, each one of these circles intersects the fixed Ewald circle at a pair of points $\mathbf{k}'$, both forming an angle 2θ with $\mathbf{k}$. In 3D, each sphere intersects the Ewald sphere at a circle of points $\mathbf{k}'$, again forming a fixed angle 2θ with $\mathbf{k}$: therefore scattered radiation comes out in coaxial *cones* of $\mathbf{k}'$ directions, axially symmetric around the $\hat{\mathbf{k}}$ direction. $\mathbf{G}$ vectors longer than $2|\mathbf{k}|$ generate circles/spheres such as the *dashed one*, which do not intersect the Ewald sphere, thus producing no diffraction

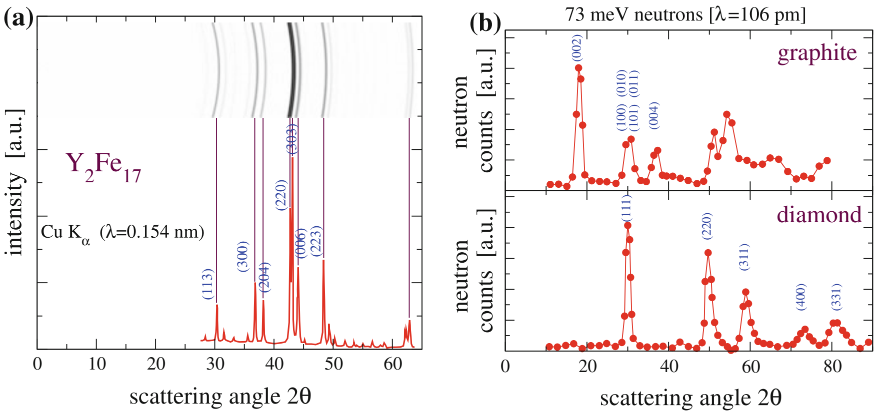


Fig. 5.40 Powder diffraction patterns: intensity as a function of the scattering angle 2θ. **a** Scattering of the copper K_α X-ray radiation by a Y_2Fe_{17} powder sample, with the intensity counting compared to the rings on the film recording. (Data from Ref. [38].) **b** Scattering of low-energy neutrons by powder samples of elementary carbon (Data from Ref. [39])

forms *cones* whose axis is the direction of the incident radiation $\hat{\mathbf{k}}$. The rings of Fig. 5.40a illustrate an example of such a pattern recorded on film. Also Fig. 5.11 shows an example of such a pattern, generated by an Al foil (not a proper powder): the sharp cones of radiation diffracted at characteristic angles prove that Al is microcrystalline. Microcrystalline structures of this kind (tightly bound collections of randomly oriented microscopic individual crystals separated by grain boundaries) are responsible for the plastic deformable character of most solid metals, as opposed to the rigidity of those solids which crystallize in the form of large single crystals (e.g. Si, NaCl).

Structural data about powder or microcrystalline samples are conveniently collected in plots of the diffracted intensity as a function of the angle 2θ, as in the patterns of Fig. 5.40. In this kind of diffractograms, the vertical axis reports the total scattered intensity, integrated along circles at fixed 2θ.

The derivation of Eq. (5.17) is readily generalized to crystals with n_d atoms in each periodically repeated cell, where the electron density is $n_{\rm el}(\mathbf{r}) = \sum_{\mathbf{R}} \sum_{j=1}^{n_d} n_{{\rm at}\,j}(\mathbf{r} - \mathbf{d}_j - \mathbf{R})$. The result is that the atomic form factor in Eqs. (5.16–5.18) must be replaced by a *structure factor*

$$S(\mathbf{q}) = \sum_{j=1}^{n_d} e^{-i\mathbf{q}\cdot\mathbf{d}_j} f_{{\rm at}\,j}(\mathbf{q}), \tag{5.23}$$

representing the Fourier transform of the matter distribution of the n_d atoms sitting at positions $\mathbf{d}_j$ in the repeated cell. Expression (5.23) holds for X-rays, $f_{{\rm at}\,j}(\mathbf{q})$ being the atomic form factors of the individual atoms in a cell. For neutrons, the same expression applies, provided that the $f_{{\rm at}\,j}(\mathbf{q})$ are replaced by $\mathbf{q}$-independent neutron scattering amplitudes of the individual nuclei in the cell. The decomposition $I(\mathbf{q}) \propto |\tilde{n}_{\rm Bravais}(\mathbf{q})|^2 \, |S(\mathbf{q})|^2$ implies a fundamental result: a given Bravais lattice yields the same characteristic diffraction pattern irrespective of the number, kinds and positions of the atoms populating its unit cell: these details affect the structure factor $|S(\mathbf{q})|^2$, in turn applying a multiplicative intensity modulation to the *same* peaks.

The possibility of describing a same crystal structure in terms of different lattices with different basis might make us worry of some ambiguity. For example, a square lattice of side a could also be viewed as a square lattice of side $2a$ with $n_d = 4$ atoms per cell: the reciprocal-lattice $\mathbf{G}$ points of the $2a$ lattice are twice as dense in each direction, but the diffracted pattern must remain the same, because this is the same crystal described in a formally different way. Indeed, the structure factor (drawn in Fig. 5.33a) vanishes for all $\mathbf{G}$ points of the denser reciprocal lattice which do not belong to the reciprocal lattice of the true a-side square. However, the $2a$ lattice may become the correct minimal description of the actual structure, e.g. as a consequence of a structural deformation—one or several atoms in each 2×2 square moving away from their perfect-square position. With such a deformation, some of the $2a$ denser $\mathbf{G}$ peaks acquire nonzero intensity. Likewise, a fcc crystal of conventional cell side a, as a Bravais lattice, generates diffraction spots for all $\mathbf{q} = \mathbf{G}$ points forming a bcc reciprocal lattice of conventional side $4\pi/a$. However, the same fcc structure can be viewed as a simple-cubic lattice of side a with 4 atoms/cell. The reciprocal lattice of this simple-cubic lattice has denser $\mathbf{G}$ points (a simple-cubic lattice of side $2\pi/a$) than that of the fcc. However, the diffraction pattern is the same, regardless of the adopted formalism, because, the structure factor (5.23), computed for 4 equal atoms [equal $f_{\mathrm{at}\,j}(\mathbf{q})$] sitting at positions $\mathbf{0}$, $(\hat{\mathbf{x}} + \hat{\mathbf{y}})a/2$, $(\hat{\mathbf{x}} + \hat{\mathbf{z}})a/2$, $(\hat{\mathbf{y}} + \hat{\mathbf{z}})a/2$, gets rid of peaks at $\mathbf{G}$ points of the simple-cubic lattice (side $2\pi/a$) not belonging to the actual bcc reciprocal lattice (side $4\pi/a$).

Defects and the finite crystal size displace some scattered intensity from the sharp Bragg peaks into a diffuse continuous background (see e.g. Fig. 5.30). If disorder increases, the intensity of this continuous background grows until, for amorphous or liquid samples, neutron or X-ray scattering does not exhibit the sharp Bragg peaks characteristic of lattice periodicity any more: the scattered intensity is then a smooth function of the angle 2θ, which provides useful statistical information about the materials' structural properties, retrievable by numerical Fourier analysis.

5.2 Electrons in Crystals

Within the adiabatic framework (Sect. 3.1), electrons move in a solid according to the electronic equation (3.2). The total electronic energy obtained by solving Eq. (3.2), added to the inter-nuclear repulsion, yields the adiabatic potential (3.10) which, in turn, determines the dynamics of the nuclei through Eq. (3.9). For a crystalline solid at low enough temperature, V_{ad} keeps the atomic configuration close to its minimum, characterized by a regular arrangement of the nuclei, basically a large chunk of an ideal crystal of the kind described in Sect. 5.1. For the moment, we neglect the nuclear kinetic energy and the movements of the ions: at equilibrium, we assume that all ions sit at the ideal crystal-structure positions (deferring to Sect. 5.3 the study of the ionic motions about these equilibrium configurations), and focus our attention

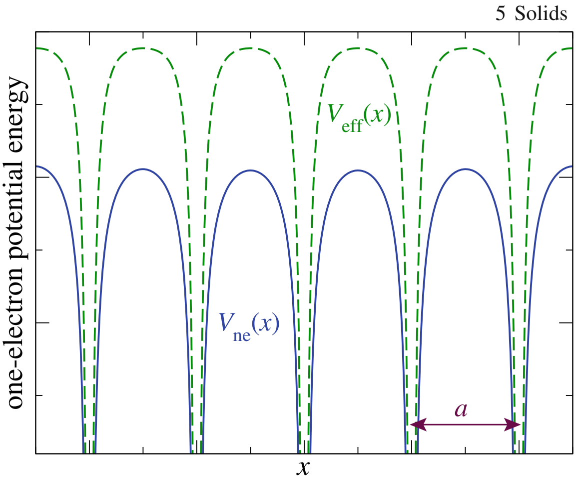


Fig. 5.41 The bare nuclear potential energy $V_{ne}(x)$ (*solid*) and the screened effective one-electron potential energy $V_{\text{eff}}(x)$ (*dashed*) experienced by an electron moving in a crystal along a line through a direct-lattice primitive direction. The screened potential energy is less attractive than that produced by the bare nuclei, but it exhibits the same lattice translational symmetry: $V_{\text{eff}}(\mathbf{r}) = V_{\text{eff}}(\mathbf{r} + \mathbf{R})$

to the motion of electrons. Within this idealized scheme, we investigate the solutions of the electronic equation (3.2), starting from their general properties required by symmetry.

The many-body equation (3.2) is plagued by the same difficulties discussed for atoms and molecules. The Schrödinger problem of many electrons in a crystal can usually be approached in a one-electron mean-field scheme, e.g. of the Hartree-Fock type, like that of atoms and molecules. As discussed in Sect. 2.2.4, this type of approximation maps the N-electron equation to single-electron self-consistent equations for the motion of one electron in the field of the nuclei and the charge distribution of the $N - 1$ other electrons. The mean-field effective potential $V_{\text{eff}}(\mathbf{r})$ has the same symmetry as the potential created by the bare nuclei, i.e. the full crystal symmetry[12] (Fig. 5.41). Similarly to atoms, where the single-electron HF orbitals carry spherical-symmetry labels l and m, the single-electron states of crystals carry Bravais-lattice group representations, labeled by $\mathbf{k}$ vectors chosen within a primitive zone of the reciprocal lattice, often the first BZ. Specifically, the $\mathbf{k}$ quantum number carries information about the way a wavefunction changes under the action of lattice translation $T_{\mathbf{R}}$, i.e. in going from one lattice cell to the next. Like in atoms, additional non-symmetry-related quantum numbers identify states of different nodal structure (within each primitive cell) and spin projection.

---

[12] This assumption is more well-grounded than the approximation of a spherically-symmetric mean field in atoms. The reason is that, while non-s atomic states have nonuniform angular distributions, as discussed below crystal states have the same probability distribution (wavefunction square modulus) in all cells. As a related example, the symmetric and antisymmetric wavefunctions of Eq. (3.11), drawn in Fig. 3.3, have both a perfectly symmetric square modulus under reflections across the mid-plane separating the nuclei.

The determination of the eigenfunction and eigenenergy for an electron moving according to $H_1 = [T_{e\,1} + V_{\text{eff}}(\mathbf{r})]$ requires a detailed calculation, which is usually carried out numerically. In this context, symmetry plays a twofold role:

- simplify greatly the solution of the HF-Schrödinger equation in the crystal;
- understand general features of its solutions.

Bloch's theorem expresses the action of lattice discrete symmetry on the electronic eigenstates in crystals. This theorem states that in a periodic potential $[V_{\text{eff}}(\mathbf{r}) = V_{\text{eff}}(\mathbf{r}+\mathbf{R})]$ all Schrödinger eigenstates can be chosen in the factorized form

$$\psi_j(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}\,j}(\mathbf{r}), \tag{5.24}$$

where the function $u_{\mathbf{k}\,j}(\mathbf{r})$ has the same periodicity of the lattice $[u_{\mathbf{k}\,j}(\mathbf{r}) = u_{\mathbf{k}\,j}(\mathbf{r}+\mathbf{R})]$, and $\mathbf{k}$ is a suitable wave vector (depending on ψ_j, but otherwise subject to no restriction). This fundamental result can be interpreted in two alternative, but equally instructive ways:

- In a periodic context all electronic eigenfunctions display a nontrivial spatial dependence within one primitive unit cell only: in any other cell displaced by $T_{\mathbf{R}}$, the wavefunction is equal to the one in the original cell, apart from a constant phase factor $e^{i\mathbf{k}\cdot\mathbf{R}}$ which leaves the probability distribution $|\psi_j|^2$ unaffected. This consideration is basically also the demonstration of Bloch's theorem.[13]
- In a periodic potential, the Schrödinger eigenstates are essentially free-electron plane-wave–like states, except for a periodic (thus trivial) amplitude modulation.

Observe that $\mathbf{k}$ may as well be restricted to one primitive cell of the reciprocal lattice, e.g. a primitive parallelepiped, or the first BZ[14]—see Fig. 5.17. Indeed, if $\mathbf{k}$ was outside this primitive cell, we could always find a reciprocal lattice vector $\mathbf{G}$ such that $\mathbf{k}' = \mathbf{k} + \mathbf{G}$ is in the primitive cell of our choice. But then $\psi_j(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}\,j}(\mathbf{r}) = e^{i(\mathbf{k}'-\mathbf{G})\cdot\mathbf{r}} u_{\mathbf{k}\,j}(\mathbf{r}) = e^{i\mathbf{k}'\cdot\mathbf{r}} e^{-i\mathbf{G}\cdot\mathbf{r}} u_{\mathbf{k}\,j}(\mathbf{r})$, and the function $u'_{\mathbf{k}'\,j}(\mathbf{r}) = e^{-i\mathbf{G}\cdot\mathbf{r}} u_{\mathbf{k}\,j}(\mathbf{r})$ is lattice-periodic, thus $e^{i\mathbf{k}'\cdot\mathbf{r}} u'_{\mathbf{k}'\,j}(\mathbf{r})$ is a valid Bloch function.

It is instructive to find the explicit equation satisfied by the Bloch functions $u_{\mathbf{k}\,j}(\mathbf{r})$. This is obtained by substituting the decomposition (5.24) in the stationary Schrödinger equation for the electrons in the periodic effective potential:

$$\left[-\frac{\hbar^2}{2m_e}\nabla^2 + V_{\text{eff}}(\mathbf{r})\right] e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) = \mathcal{E}\, e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}). \tag{5.25}$$

---

[13] The single-electron effective Hamiltonian $T_e + V_{\text{eff}}(\mathbf{r})$ commutes with all lattice translations $T_{\mathbf{R}}$. Accordingly, its eigenfunctions may be chosen as simultaneous eigenfunctions of all discrete translation operators. But then $T_{\mathbf{R}}\psi_j(\mathbf{r}) = e^{-i\mathbf{k}'\cdot\mathbf{R}}\psi_j(\mathbf{r})$ for some $\mathbf{k}'$ in the first BZ. Accordingly, one can call $\mathbf{k} = -\mathbf{k}'$ and construct $u_{\mathbf{k}\,j}(\mathbf{r}) = \psi_j(\mathbf{r})\, e^{i\mathbf{k}'\cdot\mathbf{r}}$, which is periodic, as can be readily verified by the substitution of $\mathbf{r} - \mathbf{R}$ in place of $\mathbf{r}$.

[14] In 1D, any k-interval of length $2\pi/a$. For example the first BZ, i.e. the interval $-\pi/a < k \le \pi/a$.

To deal with the kinetic term, observe that

$$\begin{aligned}\nabla^2 e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}\,j}(\mathbf{r}) &= \nabla \cdot \left[e^{i\mathbf{k}\cdot\mathbf{r}} \nabla u_{\mathbf{k}\,j}(\mathbf{r}) + i\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}\,j}(\mathbf{r})\right] \\ &= e^{i\mathbf{k}\cdot\mathbf{r}} \nabla^2 u_{\mathbf{k}\,j}(\mathbf{r}) + 2e^{i\mathbf{k}\cdot\mathbf{r}} i\mathbf{k}\cdot\nabla u_{\mathbf{k}\,j}(\mathbf{r}) - |\mathbf{k}|^2 e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}\,j}(\mathbf{r}) \\ &= e^{i\mathbf{k}\cdot\mathbf{r}} \left(\nabla + i\mathbf{k}\right)^2 u_{\mathbf{k}\,j}(\mathbf{r}).\end{aligned}$$

By substituting this decomposition into the Schrödinger equation (5.25), and dividing by the common factor $e^{i\mathbf{k}\cdot\mathbf{r}}$, we obtain

$$\left[-\frac{\hbar^2}{2m_e}\left(\nabla + i\mathbf{k}\right)^2 + V_{\text{eff}}(\mathbf{r})\right] u_{\mathbf{k}\,j}(\mathbf{r}) = \mathscr{E}_{\mathbf{k}\,j}\, u_{\mathbf{k}\,j}(\mathbf{r}). \tag{5.26}$$

This fundamental equation allows us to compute the stationary states of every electron in a crystal, given its wave number $\mathbf{k}$. Thanks to the periodicity of $u_{\mathbf{k}\,j}$ established by Bloch's theorem, Eq. (5.26) must be solved within a *single cell* of the direct lattice (with applied periodic boundary conditions) rather than over the macroscopically large volume of the *whole crystal*. In practice, the need to solve Eq. (5.26) in a microscopically small unit cell represents an immense technical advantage. No such possibility is available in the absence of lattice periodicity, as in non-periodic solids, where the electronic motions are therefore far harder to investigate. Once $u_{\mathbf{k}\,j}(\mathbf{r})$ is evaluated by exact or approximate techniques, the full electronic wavefunction $\psi_j(\mathbf{r})$ can then be extended to the whole crystal by means of Eq. (5.24).

Before attempting the solution of Eq. (5.26), we discuss the qualitative properties of the single-electron eigenstates and eigenenergies in a crystal. Equation (5.26) is a second-order differential equation, defined in the unit-cell volume V_c, with standard periodic boundary conditions. Apart for the wave-vector shift by $\mathbf{k}$, Eq. (5.26) is equivalent to a stationary Schrödinger equation. Accordingly, for any fixed $\mathbf{k}$, its solutions must be analogous to those of a standard Schrödinger equation (B.30) in a microscopically small volume, namely: a ladder of discrete eigenenergies $\mathscr{E}_{\mathbf{k}\,j}$ associated to eigenfunctions $u_{\mathbf{k}\,j}(\mathbf{r})$, characterized by an increasing number of nodal planes for increasing energy. The index $j = 1, 2, 3, \ldots$ labels precisely these solutions, for increasing energy, see Fig. 5.42.

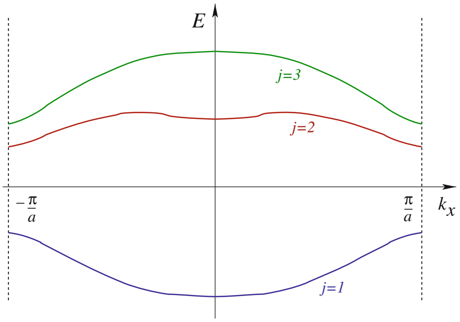


Fig. 5.42 For each given $\mathbf{k}$ in the first BZ, the solution of Eq. (5.26) provides discrete electronic energies $\mathscr{E}_{\mathbf{k}\,j}$, for $j = 1, 2, \ldots$, three of which are sketched here. These energies depend parametrically on $\mathbf{k}$. The continuous functional dependence $\mathscr{E}_j(\mathbf{k}) = \mathscr{E}_{\mathbf{k}\,j}$ is an energy band of the solid

$\mathbf{k}$ can vary arbitrarily within the first BZ, and as it does, the Eq. (5.26) for $u_{\mathbf{k}}(\mathbf{r})$ changes. This justifies attaching the label $\mathbf{k}$ to eigenenergies $\mathcal{E}_{\mathbf{k}\,j}$ and eigenfunctions $u_{\mathbf{k}\,j}(\mathbf{r})$. The parameter $\mathbf{k}$ affects Eq. (5.26) analytically, thus one should expect that its solutions depend analytically on $\mathbf{k}$. Indeed, for fixed j, the eigenenergies $\mathcal{E}_{\mathbf{k}\,j}$ depend on $\mathbf{k}$ as continuous functions called *energy bands*, or simply "bands" (Fig. 5.42). As $\mathbf{k}$ takes all its allowed values within the first BZ, for each j, $\mathcal{E}_{\mathbf{k}\,j}$ spans a continuous interval of available energies (the range of the $\mathcal{E}_{\mathbf{k}\,j}$ function, sometimes itself called a "band"). The ranges of two successive bands $\mathcal{E}_{\mathbf{k}\,j}$ and $\mathcal{E}_{\mathbf{k}\,j+1}$ can either overlap (like bands 2 and 3 in Fig. 5.42) or not overlap (like bands 1 and 2 in Fig. 5.42). Both possibilities are compatible with Eq. (5.26), and do occur in actual solids. The main consequence of Bloch's theorem is then a spectrum of electronic energies involving intervals of allowed energies, often separated by ranges of forbidden energies (*band gaps*). In a crystal, electrons are characterized by an energy spectrum somewhat intermediate between that of a free particle (all positive energies) and that of an atom (isolated eigenvalues separated by gaps).

Figure 5.43 shows the qualitative form of the band eigenfunctions in a 10-sites portion of a monoatomic 1D crystal (this sketch could be extended to 2D and 3D). The real part of the wavefunctions is drawn for two bands: the imaginary part is similar, although with a different phase, so that $|\psi_{\mathbf{k}\,j}(x)|^2$ is periodic, i.e. repeated equally in all cells. The primary role of $\mathbf{k}$ is to tune the phase change of the wavefunction in going from one site to the next. The wavelength associated to these *Bloch waves* is $2\pi/|\mathbf{k}|$. The values of $\mathbf{k}$ selected in Fig. 5.43 yield wavelengths commensurate to the 10-sites region drawn. Other intermediate choices of $\mathbf{k}$ would produce wavelengths

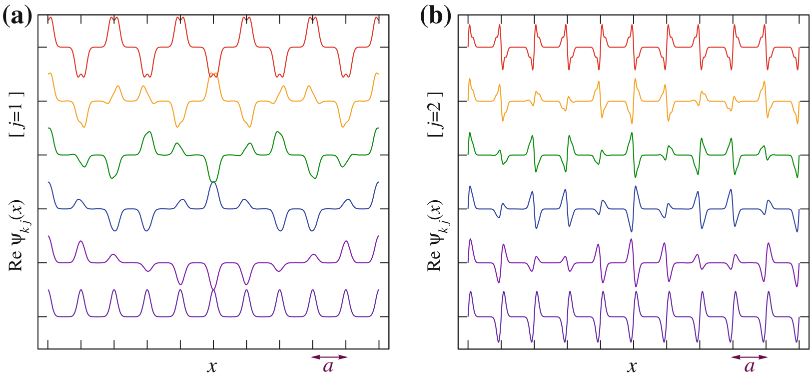


Fig. 5.43 A sketch of the real part of a few Bloch wavefunction $\psi_{k\,j}(x) = e^{ik\cdot x}\,u_{k\,j}(x)$ as a function of position x in a 1D crystal, for (*bottom* to *top*) $k = 0$, $k = 0.1 \times 2\pi/a$, $k = 0.2 \times 2\pi/a$, $k = 0.3 \times 2\pi/a$, $k = 0.4 \times 2\pi/a$, $k = 0.5 \times 2\pi/a$, for the states belonging to bands **a** $j = 1$ and **b** $j = 2$. Negative k values of the same length $|k|$ yield wavefunctions whose real part is the same. As k increases from 0 to the first-BZ boundary, the phase difference of the wavefunction at neighboring sites increases. At the zone boundary $k = \pi/a$, this phase difference is maximum, and equals π, corresponding to the sign alternation of the *topmost curve*. At the same time, due to the k dependence of Eq. (5.26) for $u_{k\,j}(x)$, the shape of the wavefunction in each cell also evolves with k

not matching the boundary condition. In addition, **k** affects also the local "shape" of the wavefunction $\psi_{\mathbf{k}\,j}(x)$ within each cell, through the explicit dependence of Eq. (5.26), and thus of $u_{\mathbf{k}\,j}(x)$. This **k**-dependence is smooth and often relatively mild, while the wavefunctions usually change more when comparing different bands j and j'.

The detailed calculation of the band energies $\mathscr{E}_{\mathbf{k}\,j}$ and wavefunctions $u_{\mathbf{k}\,j}(x)$ of an actual solid is typically carried out by means of some self-consistent calculation of the mean-field potential $V_{\mathrm{eff}}(\mathbf{r})$, associated to the numerical solution of Eq. (5.26) in a direct-lattice primitive unit cell, for a sufficiently dense set of **k** points. However, useful insight in the physics of the band states can be obtained by considering substantially simplified "model" solutions of Eq. (5.26).

5.2.1 Models of Bands in Crystals

Although for many crystals Eq. (5.26) is not especially difficult to solve numerically in a primitive unit cell, for any reasonable number of "sample" **k** points, approximate solutions help us to gain insight in the fundamental properties of the band energies and states. Like in Sect. 4.3.1, we apply periodic boundary conditions to a macroscopic box, chosen of a shape and volume multiple of V_c, suitable to preserve the lattice symmetry. We refrain from applying more realistic open-end boundary conditions, which would break the periodicity, and would involve the subtle properties of *crystal surfaces*, themselves an entire branch of science.

5.2.1.1 The Tight-Binding Model

In a region near each atomic nucleus, the crystal effective potential (Fig. 5.41) acting on an electron resembles that of an isolated atom. Accordingly, for the crystal electronic wavefunctions it makes sense to follow the construction of molecular orbitals in terms of atomic ones, as sketched in Chap. 3 for H_2^+ and H_2. In this picture, a crystal is viewed as a huge molecule with the band states constructed as the natural generalization of bonding and antibonding molecular orbitals. Models for extended electronic states expressed as *linear combinations of atomic orbitals* are often referred to as "tight binding".

The 2-atoms calculation of Sect. 3.2.1 is the simplest example of tight-binding application. It addresses the 1s "band" of a "crystal" composed of $N_n = 2$ H atoms only. $|1sL\rangle$ represents the state at the left atom. $|1sR\rangle$, represents that at the right atom, obtained from $|1sL\rangle$ by a "lattice" translation. Fictitious periodic boundary conditions bring $|1sR\rangle$ back again to $|1sL\rangle$ upon a further translation (Fig. 5.44). The symmetry-adapted states of Eq. (3.11) can be viewed as bonding and antibonding combinations

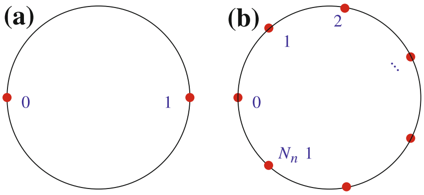


Fig. 5.44 **a** A tiny “crystal” of $N_n = 2$ equal atoms represents a diatomic molecule (with its bond duplicated). **b** A 1D crystal of N_n equal atoms with periodic boundary conditions exhibits the connectivity of a ring

$$|S\rangle = \mathscr{N}_S(|1sL\rangle + |1sR\rangle) = \mathscr{N}_S \sum_{p=0,1} |p\rangle = \mathscr{N}_S \sum_{p=0}^{N_n-1} e^{ikpa}|p\rangle, \quad k = 0$$

$$|A\rangle = \mathscr{N}_A(|1sL\rangle - |1sR\rangle) = \mathscr{N}_A \sum_{p=0,1} (-1)^p\,|p\rangle = \mathscr{N}_A \sum_{p=0}^{N_n-1} e^{ikpa}|p\rangle, \quad k = \frac{\pi}{a},$$

with $|0\rangle = |1sL\rangle$, $|1\rangle = |1sR\rangle$, and $\mathscr{N}_{S/A}$ are suitable normalization constants. The only possible phase relations in these two-site combinations are “in phase” and “out of phase”.

These linear combinations can be generalized to a chain of $N_n \geq 2$ hydrogen atoms, forming a 1D “crystal”, Fig. 5.44. To construct kets compatible with the global “ring” periodic boundary condition $\langle r + N_n a|\phi\rangle = \langle r|\phi\rangle$, 1D version of Eq. (B.55) with $L = N_n a$, we adopt the phase factors generated by:

$$k = \left(-\frac{1}{2} + \frac{1}{N_n}\right)\frac{2\pi}{a},\ \left(-\frac{1}{2} + \frac{2}{N_n}\right)\frac{2\pi}{a},\ \left(-\frac{1}{2} + \frac{3}{N_n}\right)\frac{2\pi}{a}, \ldots,\ \frac{1}{2}\,\frac{2\pi}{a}. \quad (5.27)$$

We construct N_n combinations of the 1s orbitals by using the appropriate phase factors in the linear combination:

$$|\phi_k\rangle = \mathscr{N}_k \sum_{p=0}^{N_n-1} e^{ikpa}|p\rangle. \quad (5.28)$$

In the limit of macroscopically large N_n, the energies of the $|\phi_k\rangle$ states form a continuum in the first BZ $-\pi/a < k \leq \pi/a$. As k spans the BZ, the band energy spans continuously the range from the bonding value (band bottom) to the antibonding value (band top), Fig. 5.45.

By extending the derivation of the molecular Eq. (3.13), we can work out a method to compute the band energies for a general crystal, which will provide an approximate relation for the 1D band of Fig. 5.45 as a special case. The tight-binding method is based on expanding the crystal electronic states as linear combinations of eigenstates of the atomic one-electron effective Hamiltonian H^{at}:

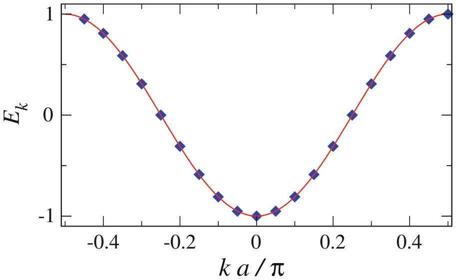


Fig. 5.45 As the N_n equally-spaced k points in the first BZ allowed by the entire-sample artificial periodicity become denser and denser in the $N_n \to \infty$ macroscopic-crystal limit, the electronic energies form the band continuum of a macroscopic solid

$$|\psi\rangle = \sum_{\mathbf{R}} \sum_{n} b_{\mathbf{R}n} T_{\mathbf{R}} \, |\phi_n\rangle. \tag{5.29}$$

In this expression, appropriate for a monoatomic crystal, $|\phi_n\rangle$ is the nth single-electron eigenstate of the isolated atom, $T_{\mathbf{R}}$ is an operator which translates rigidly by a lattice translation $\mathbf{R}$, and $b_{\mathbf{R}\,n}$ are suitable coefficients. One usually selects a subset $n = 1, \ldots, n_o$ of the atomic states, usually comprising a few valence orbitals only, and solves the one-electron Hamiltonian problem (B.36) in its variational matrix form, see Appendix B.5. The involved matrix elements of the 1-electron Hamiltonian are $H_{mn}(\mathbf{R}) = \langle\phi_m|H_1\, T_{\mathbf{R}}|\phi_n\rangle$, and the overlap integrals are $B_{mn}(\mathbf{R}) = \langle\phi_m|T_{\mathbf{R}}|\phi_n\rangle$. The construction of the eigenstates and eigenvalues of Eq. (B.36) involves the (generalized) diagonalization of a square matrix whose size equals the number n_o of orbitals involved in each unit cell multiplied by the number of unit cells in the crystal. This prohibitively large size can be reduced to just n_o, by taking advantage of the discrete translational symmetry.

We decompose the one-electron Hamiltonian $H_1 = [T_{e\,1} + V_{\text{eff}}(\mathbf{r})]$ as

$$H_1 = H^{\text{at}} + \Delta U, \tag{5.30}$$

where H^{at} contains the kinetic energy $T_{e\,1}$ plus the effective potential energy for an electron moving around one *isolated* ion placed at the origin $\mathbf{R} = \mathbf{0}$, and $\Delta U(\mathbf{r})$ is the total screened potential energy generated by all ions in the crystal, reduced by the single-ion contribution, see Fig. 5.46. In terms of the atomic levels E_n^{at} and matrix elements

$$C_{mn}(\mathbf{R}) = \langle\phi_m|\Delta U\; T_{\mathbf{R}}\,|\phi_n\rangle, \tag{5.31}$$

the tight-binding matrix secular problem (B.36) can be expressed as

$$\sum_{n\,\mathbf{R}} \left[E_m^{\text{at}} B_{mn}(\mathbf{R}) + C_{mn}(\mathbf{R})\right] b_{\mathbf{R}\,n} = \mathscr{E} \sum_{n\,\mathbf{R}} B_{mn}(\mathbf{R})\, b_{\mathbf{R}\,n}. \tag{5.32}$$

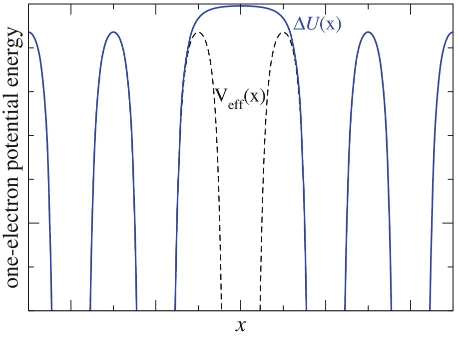


Fig. 5.46 The screened one-electron potential energy $V_{\text{eff}}(x)$ (*dashed*) experienced by an electron moving in the crystal, and the same potential to which the contribution of the ion at $\mathbf{R} = \mathbf{0}$ has been subtracted ($\Delta U(x)$, *solid*). The plot sketches a 1D cut along the x axis of a typical crystal

Inspired by Eqs. (5.24) and (5.28), we guess an appropriate **R**-dependence of the coefficients $b_{\mathbf{R}\,n}$, to reconstruct the appropriate Bloch symmetry:

$$b_{\mathbf{R}\,n} = e^{i\mathbf{k}\cdot\mathbf{R}}\,\tilde{b}_n. \tag{5.33}$$

In practice, Eq. (5.33) imposes a specific amplitude and inter-cell phase relation to the wavefunction $\langle \mathbf{r}|\psi\rangle$, as dictated by the selected **k**. The wavefunction $\langle \mathbf{r}|\psi\rangle$ satisfies Bloch's theorem, as is readily verified. Eigenstates are obtained by selecting appropriate combinations, as dictated by the coefficients $\tilde{b}_n$.

For given **k**, we indicate by $|\mathcal{E}_{\mathbf{k}}\rangle$ the Bloch wavefunction $|\psi\rangle$ provided that it solves the Schrödinger equation

$$H_1|\mathcal{E}_{\mathbf{k}}\rangle = \mathcal{E}_{\mathbf{k}}|\mathcal{E}_{\mathbf{k}}\rangle, \tag{5.34}$$

where the eigenvalue $\mathcal{E}_{\mathbf{k}}$ is the band eigenenergy. By multiplication on the left by $\langle \phi_m|$, Eq. (5.34) for $|\mathcal{E}_{\mathbf{k}}\rangle$ maps to a matrix equation for the expansion coefficients $\tilde{b}_n$:

$$\sum_n H_{\mathbf{k},mn}\tilde{b}_n = \mathcal{E}_{\mathbf{k}} \sum_n B_{\mathbf{k},mn}\,\tilde{b}_n. \tag{5.35}$$

For each fixed **k**, Eq. (5.35) is the generalized eigenvalue problem for the $n_o \times n_o$ matrix $H_{\mathbf{k}}$, of the type of Eq. (B.36). The n_o eigenvalues obtained for each **k** point form n_o bands, as **k** is let vary across the BZ. In practice the matrix diagonalization (5.36) needs to be carried out at a sufficiently fine discrete mesh of **k** points. $\mathcal{E}_{\mathbf{k}}$ represents the generic eigenvalue, corresponding to an eigenvector $\tilde{b}$ of components $\tilde{b}_n$. The solution eigen-coefficients $\tilde{b}_n$ depend on **k** and on the precise eigenvalue (band) considered. In the matrix multiplication language, Eq. (5.35) reads simply

$$H_{\mathbf{k}} \cdot \tilde{b} = \mathcal{E}_{\mathbf{k}}\; B_{\mathbf{k}} \cdot \tilde{b}. \tag{5.36}$$

The tight-binding energy and overlap matrices of Eq. (5.35) are explicitly:

$$H_{\mathbf{k},mn} = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \langle \phi_m | H_1 \, T_{\mathbf{R}} | \phi_n \rangle \tag{5.37}$$

$$B_{\mathbf{k},nm} = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \langle \phi_m | \, T_{\mathbf{R}} \, | \phi_n \rangle, \tag{5.38}$$

with the sums over $\mathbf{R}$ extending in principle over all lattice points. In practice, when the distance $|\mathbf{R}|$ grows significantly beyond the typical atomic size (a few times a_0), the energy and overlap matrix elements, decay rapidly to 0, due to the exponential decay of the atomic wavefunctions $\langle \mathbf{r} | \phi_n \rangle$ at large distance $|\mathbf{r}|$.

Using the decomposition (5.30), Eq. (5.35) can be written in the instructive form related to Eq. (5.32):

$$\sum_n \left[E_n^{\text{at}} B_{\mathbf{k},mn} + C_{\mathbf{k},mn} \right] \tilde{b}_n = \mathscr{E}_{\mathbf{k}} \sum_n B_{\mathbf{k},mn} \tilde{b}_n, \tag{5.39}$$

where the matrix $C_{\mathbf{k}}$ is the Fourier transform of Eq. (5.31):

$$C_{\mathbf{k},mn} = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \langle \phi_m | \Delta U \, T_{\mathbf{R}} | \phi_n \rangle. \tag{5.40}$$

Assuming, for simplicity, that the overlap matrix is the identity ($B_{\mathbf{k},mn} \approx \delta_{mn}$), Eq. (5.39) takes the following explicit form:

$$\begin{pmatrix} E_1^{\text{at}} + C_{\mathbf{k},11} & C_{\mathbf{k},12} & C_{\mathbf{k},13} & \cdots \\ C_{\mathbf{k},21} & E_2^{\text{at}} + C_{\mathbf{k},22} & C_{\mathbf{k},23} & \cdots \\ C_{\mathbf{k},31} & C_{\mathbf{k},32} & E_3^{\text{at}} + C_{\mathbf{k},33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \cdot \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \tilde{b}_3 \\ \vdots \end{pmatrix} = \mathscr{E}_{\mathbf{k}} \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \tilde{b}_3 \\ \vdots \end{pmatrix}. \tag{5.41}$$

Figure 5.47 sketches the connection between the atomic and the solid-state spectra as the component atoms, sodium in this example, move together to form a crystal. Equation (5.41) provides a quantitative justification for the physics of Fig. 5.47: for very large interatomic separations, all $\mathbf{k}$-dependent part of the matrix elements $C_{\mathbf{k},mn}$ vanishes exponentially and the overlap matrix $B_{\mathbf{k},mn}$ does coincide with δ_{mn}, thus the eigenvalue problem (5.41) becomes trivially independent of $\mathbf{k}$ and the bands become flat and coinciding with the atomic eigenenergies E_m^{at}. As the atoms move closer together, $|C_{\mathbf{k},mn}|$ generally increase, thus leading to proper band dispersions with finite *bandwidth* (i.e. the difference $\mathscr{E}_{\text{max}} - \mathscr{E}_{\text{min}}$). Further shrinking of the lattice periodicity (as induced, e.g., by the application of hydrostatic pressure to the crystal), tends to broaden the band dispersion and a consequently decrease the forbidden energy gaps. Similarly, for a given interatomic separation, since high up excited atomic states decay relatively slowly with distance (see, e.g., Fig. 2.6), these

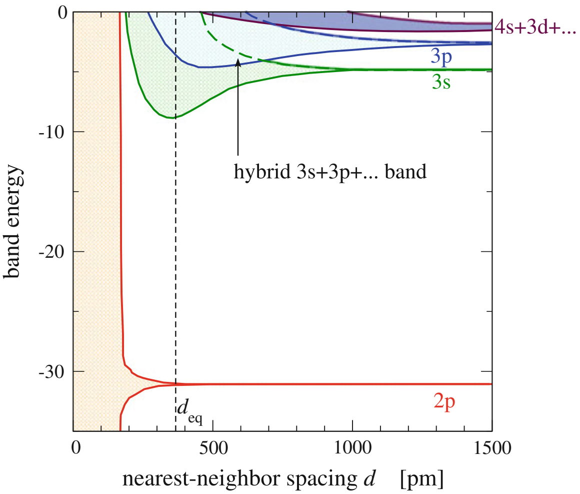


Fig. 5.47 The energy bands of bcc sodium, as a function of the nearest-neighbor interatomic spacing $d = a\sqrt{3}/2$. As the atoms move closer, their separation becomes comparable to the size of the atomic wavefunctions, and the individual atomic levels spread out into bands. If the lattice parameter decreases, e.g. under the effect of an external pressure, then all bandwidths tend to increase further, and more bands may overlap. At the equilibrium spacing $d_{\rm eq} = 367\,\text{pm}$, the strong overlap of the 3s, 3p,… bands indicates that the tight-binding method is not especially well suited to describe this wide highly hybridized conduction band. Indeed, the plane-waves method (Sect. 5.2.1.2) describes the conduction band of alkali metals better

states give origin to generally larger $|C(\mathbf{k})_{mn}|$, thus to broader band dispersions than the more localized deep (core) levels.

The tight-binding method, requiring a number of (small) matrix diagonalizations, is easily coded in a computer program, and it is employed routinely as a simple technique to obtain semi-quantitative band structures with a modest effort.

In the extreme limit where only $n_{\rm o} = 1$ atomic s state $|\phi_n\rangle$ is retained, the tight-binding model is conveniently solved analytically, as matrices turn into numbers, and matrix products turn into simple products. The "eigenvector" can be taken $\tilde{b}_n = 1$, and Eq. (5.39) becomes

$$E_n^{\rm at} B_{\mathbf{k},nn} + C_{\mathbf{k},nn} = \mathscr{E}_{\mathbf{k}} B_{\mathbf{k},nn}, \tag{5.42}$$

with solution

$$\mathscr{E}_{\mathbf{k}\,n} = E_n^{\rm at} + \frac{C_{\mathbf{k},nn}}{B_{\mathbf{k},nn}}. \tag{5.43}$$

We retain the label n to $\mathscr{E}_{\mathbf{k}}$ to remind us that it represents the band of atomic nature $|\phi_n\rangle$.

To further simplify Eq. (5.43), in the Fourier sums (5.38) and (5.40) of the $B_{\mathbf{k}}$ and $C_{\mathbf{k}}$ matrices, it is convenient to separate the $\mathbf{R} = \mathbf{0}$ contribution from $\mathbf{R} \neq \mathbf{0}$ terms.

The $\mathbf{R} = \mathbf{0}$ contribution to $B_{\mathbf{k}}$ equals unity, and that to $C_{\mathbf{k}}$ is a negative energy

$$\alpha_n = \langle \phi_n | \Delta U | \phi_n \rangle. \tag{5.44}$$

For the $\mathbf{R} \neq \mathbf{0}$ integrals contributing to $B_{\mathbf{k}}$ and $C_{\mathbf{k}}$ in (5.38) and (5.40) we use the notation:

$$\beta_n(\mathbf{R}) = \langle \phi_n | T_{\mathbf{R}} | \phi_n \rangle \tag{5.45}$$

$$\gamma_n(\mathbf{R}) = \langle \phi_n | \Delta U \, T_{\mathbf{R}} | \phi_n \rangle. \tag{5.46}$$

Observe that (i) by symmetry $\beta_n(-\mathbf{R}) = \beta_n(\mathbf{R})$, $\gamma_n(-\mathbf{R}) = \gamma_n(\mathbf{R})$, and (ii) for arbitrary Bravais lattice, when some $\mathbf{R}$ occurs in the $\mathbf{R}$-sums, also $-\mathbf{R}$ is present. One can then replace the complex exponentials with cosines, and rearrange the solution (5.43) for $\mathscr{E}_{\mathbf{k}\,n}$ as:

$$\mathscr{E}_{\mathbf{k}\,n} = E_n^{\text{at}} + \frac{\alpha_n + \sum_{\mathbf{R}\neq 0} \cos(\mathbf{k}\cdot\mathbf{R})\, \gamma_n(\mathbf{R})}{1 + \sum_{\mathbf{R}\neq 0} \cos(\mathbf{k}\cdot\mathbf{R})\, \beta_n(\mathbf{R})}. \tag{5.47}$$

This relation expresses the tight-binding band energy as an explicit function of $\mathbf{k}$, in terms of a set of numerical parameters E_n^{at}, α_n, $\gamma_n(\mathbf{R})$, and $\beta_n(\mathbf{R})$. Equations (5.44), (5.45) and (5.46) provide the explicit recipe to compute these parameters in terms of overlap integrals of atomic wavefunctions. Note in particular that the "hopping energy" $\gamma_n(\mathbf{R}) < 0$, because the phase of the overlapping long-distance tail of the s atomic wavefunction is the same for both wavefunctions in Eq. (5.46), and the potential ΔU is attractive.

As observed above, due to the localization of these atomic wavefunctions, both $\gamma_n(\mathbf{R})$ and $\beta_n(\mathbf{R})$ become exponentially small for large $|\mathbf{R}|$. It therefore makes sense to ignore all the integrals for $|\mathbf{R}| > R_{\text{max}}$, which would bring in only negligible corrections to the band structure $\mathscr{E}_{\mathbf{k}\,n}$. In particular, the simplest approximation, providing a band structure depending on a minimal number of parameters, is to neglect *all* of the $\beta_n(\mathbf{R})$'s (so that the denominator $B_{\mathbf{k},nn}$ becomes unity) and to include only the $\gamma_n(\mathbf{R})$'s for the nearest neighbors (nn) of $\mathbf{0}$. In this extreme simplification, the expression (5.47) becomes

$$\mathscr{E}_{\mathbf{k}\,n} = E_n^{\text{at}} + \alpha_n + \gamma_n \sum_{\mathbf{R}\in(\text{nn})} \cos(\mathbf{k}\cdot\mathbf{R}), \tag{5.48}$$

where $\gamma_n = \gamma_n(\mathbf{R}_{\text{nn}})$.

Example 1 For a 1D chain, Eq. (5.48) yields a cosine-shaped band $\mathscr{E}_k = E^{\text{at}} + \alpha + 2\gamma\,\cos(a\,k)$. Since $\gamma < 0$, the minimum of the tight-binding band is reached at $\mathbf{k} = \mathbf{0}$, as sketched in Figs. 5.45 and 5.48.[15] The molecular orbitals of the benzene

---

[15] For p orbitals the angular dependence of the atomic wavefunction can make some of the relevant $\gamma_n(\mathbf{R})$ positive, thus producing "reversed bands", with a maximum at $\mathbf{k} = \mathbf{0}$, rather than a minimum.

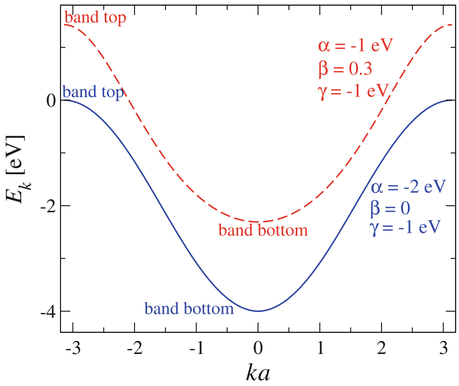


Fig. 5.48 *Solid curve* the characteristic cosine-like shape of the nearest-neighbor orthogonal 1D tight-binding band of s electrons, Eq. (5.48). *Dashed curve* the modest effect of a robust superposition integral $\beta = 0.3$, Eq. (5.47) on the band profile

ring of 6 atoms (Fig. 3.8e) can be described approximately by a 6-sites 1D-lattice tight binding, with 6 **k** points, see Eq. (5.27).

Example 2 In 2D, a square lattice has $\mathcal{E}_{\mathbf{k}} = E^{\text{at}} + \alpha + 2\,\gamma\left[\cos\left(ak_x\right) + \cos\left(ak_y\right)\right]$. Different expressions can be obtained for 2D lattices of different symmetries.

Example 3 For the fcc lattice Eq. (5.48) yields

$$\begin{aligned}\mathcal{E}_{\mathbf{k}} = E^{\text{at}} + \alpha + 2\,\gamma\Big\{ &\cos\left[\frac{a}{2}(k_x + k_y)\right] + \cos\left[\frac{a}{2}(k_y + k_z)\right] + \cos\left[\frac{a}{2}(k_z + k_x)\right] \\ &+ \cos\left[\frac{a}{2}(k_x - k_y)\right] + \cos\left[\frac{a}{2}(k_y - k_z)\right] + \cos\left[\frac{a}{2}(k_z - k_x)\right]\Big\} \\ = E^{\text{at}} + \alpha + 4\,\gamma & \\ \times \Big[\cos\left(\frac{a}{2}k_x\right)\cos\left(\frac{a}{2}k_y\right) &+ \cos\left(\frac{a}{2}k_y\right)\cos\left(\frac{a}{2}k_z\right) + \cos\left(\frac{a}{2}k_z\right)\cos\left(\frac{a}{2}k_x\right)\Big].\end{aligned}$$

The above examples of approximate band dispersions confirm that the bandwidth increases proportionally to the off-diagonal hopping energies $|\gamma|$. Changes in the overlap integral (here β) at the denominator produce only quantitative changes, but do not win over the dominant increase with $|\gamma|$. Observe also that the negative value of α reflects the fact that, on average, the energy of a band state in a crystal is lower than the original state in the isolated atom, due to the extra attraction $\Delta U(\mathbf{r})$ of all other nuclei. If the considered band happens to be filled with electrons, this energy lowering contributes to the crystal cohesive energy.

The one-orbital approximation discussed here is suitable for s bands, whenever the diagonal energy separation $\left|E_n^{\text{at}} - E_m^{\text{at}}\right|$ of $|\phi_n\rangle$ from all other atomic orbitals $|\phi_m\rangle$ is much larger than the dispersive part of $C_{\mathbf{k},mn}$. Under such conditions, the $|\phi_n\rangle$-derived s band remains isolated from all others. As soon as the inter-band gaps become comparable to the bandwidths and inter-band couplings, one needs to run a multi-orbital tight binding ($n_{\text{o}} > 1$). The bands obtained by a multi-band tight binding can be rather intricate, but they share many qualitative features with the single-band model. In particular it holds true that large interatomic separation calls for narrow bands separated by substantial gaps, and that bands get broader and more hybridized when the lattice is squeezed, as sketched in Fig. 5.47. It is apparent that for inner s shells the single-band tight-binding approximation is accurate, while it provides a rather poor description of the usually broad *conduction band* around or above the Fermi level, at the equilibrium lattice structure. An increase of the number n_{o} of the included orbitals improves systematically the accuracy of the tight-binding model, at the expense of bigger matrices to be diagonalized. Also in polyatomic crystals, the size of the matrix problem (and correspondingly the number of resulting bands) equals the total number of orbitals contributed by all atoms in a unit cell. For further details related to crystal structures with several-atom basis, refer to Ref. [10].

5.2.1.2 The Plane-Waves Method

A practical alternative to tight binding method, is the plane-waves basis. We note in Fig. 5.41 that inside a crystal the effective one-electron potential for the valence electrons is significantly screened, thus free-electron eigenstates should approximate well the valence-band eigenstates, except near the atomic nuclei. In the general discussion below, we use a 3D formalism, but this problem can be visualized more easily in 1D, where we restrict most of our illustrative examples.

Following the standard mapping of Appendix B.5, we expand the band states in the basis of momentum eigenstates:

$$|b\rangle = \sum_{\mathbf{k}'} b_{\mathbf{k}'}\, |\mathbf{k}'\rangle,$$

where $|\mathbf{k}'\rangle$ are plane-wave states, see Eqs. (B.53) and (B.56). This sum extends over all $\mathbf{k}'$ vectors, and represents therefore an integration.

Application of the Schrödinger equation to the candidate eigenket $|b\rangle$ and multiplication on the left by $\langle\mathbf{k}|$ (implying a volume integration over $\mathbf{r}$) maps the initial differential problem to an algebraic (matrix) equation for the wavefunction Fourier components $b_{\mathbf{k}'}$:

$$\begin{aligned} H|b\rangle &= \left[\frac{p^2}{2m_e} + V_{\mathrm{eff}}\right]|b\rangle = \mathscr{E}|b\rangle \\ \langle\mathbf{k}|\left[\frac{p^2}{2m_e} + V_{\mathrm{eff}}\right]|b\rangle &= \mathscr{E}\langle\mathbf{k}|b\rangle \\ \sum_{\mathbf{k}'}\left[\frac{\hbar^2 k^2}{2m_e}\delta_{\mathbf{k},\mathbf{k}'} + \langle\mathbf{k}|V_{\mathrm{eff}}|\mathbf{k}'\rangle\right] b_{\mathbf{k}'} &= \mathscr{E}\sum_{\mathbf{k}'}\delta_{\mathbf{k},\mathbf{k}'} b_{\mathbf{k}'}, \end{aligned} \tag{5.49}$$

where we use appropriate orthonormality $\langle\mathbf{k}|\mathbf{k}'\rangle = \delta_{\mathbf{k},\mathbf{k}'}$ of the plane waves and that $\mathbf{p}\,|\mathbf{k}'\rangle = \hbar\mathbf{k}'|\mathbf{k}'\rangle$. The matrix elements

$$\langle \mathbf{k}|V_{\text{eff}}|\mathbf{k}'\rangle = \mathscr{N} \int e^{i(\mathbf{k}'-\mathbf{k})\mathbf{r}} V_{\text{eff}}(\mathbf{r}) d^3\mathbf{r} = \tilde{V}_{\text{eff}}(\mathbf{k}-\mathbf{k}')$$

are the Fourier components of the potential, and $\mathscr{N}$ is the appropriate normalization, see Eqs. (B.53) and (B.56).

Until this point, we made no mention of any lattice symmetry: indeed Eq. (5.49) is the standard momentum representation of the Schrödinger equation. For a generic potential V_{eff} this formulation is not especially advantageous, since the matrix indexes $\mathbf{k}$ of the eigenvalue problem (5.49) are smooth quantities, taking continuously many values, exactly like $\mathbf{r}$ in the real-space equation. For a *periodic potential* however, as observed in the general discussion of Eq. (5.11), the Fourier expansion of the periodic potential is a *discrete Fourier series* over the reciprocal lattice, i.e. $\tilde{V}_{\text{eff}}(\mathbf{k}-\mathbf{k}')$ is nonzero only for $\mathbf{k}-\mathbf{k}' = \mathbf{G}$, a vector of the reciprocal lattice. This means that in the continuous-indexed energy matrix of Eq. (5.49) most off-diagonal matrix elements vanish. In practice, given any $\mathbf{k}'$ in the first BZ, the off-diagonal potential matrix elements connect the plane wave $|\mathbf{k}'\rangle$ only to plane waves $|\mathbf{k}\rangle$, whose $\mathbf{k}$ is displaced by a reciprocal-lattice vector: $\mathbf{k} = \mathbf{k}' + \mathbf{G}$. One can then consider separately each subset of states originated from a given $\mathbf{k}$. Only the corresponding matrix sub-blocks need to be diagonalized. In one of such sub-blocks, the matrix form of Eq. (5.49) is:

$$\begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \\ \cdots & T(\mathbf{k}+\mathbf{G}_1)+\tilde{V}_{\text{eff}}(\mathbf{0}) & \tilde{V}_{\text{eff}}(\mathbf{G}_1-\mathbf{G}_2) & \tilde{V}_{\text{eff}}(\mathbf{G}_1-\mathbf{G}_3) & \cdots \\ \cdots & \tilde{V}_{\text{eff}}(\mathbf{G}_2-\mathbf{G}_1) & T(\mathbf{k}+\mathbf{G}_2)+\tilde{V}_{\text{eff}}(\mathbf{0}) & \tilde{V}_{\text{eff}}(\mathbf{G}_2-\mathbf{G}_3) & \cdots \\ \cdots & \tilde{V}_{\text{eff}}(\mathbf{G}_3-\mathbf{G}_1) & \tilde{V}_{\text{eff}}(\mathbf{G}_3-\mathbf{G}_2) & T(\mathbf{k}+\mathbf{G}_3)+\tilde{V}_{\text{eff}}(\mathbf{0}) & \cdots \\ & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \cdot \begin{pmatrix} \vdots \\ b_{\mathbf{k}+\mathbf{G}_1} \\ b_{\mathbf{k}+\mathbf{G}_2} \\ b_{\mathbf{k}+\mathbf{G}_3} \\ \vdots \end{pmatrix}$$

$$= \mathscr{E}_{\mathbf{k}} \begin{pmatrix} \vdots \\ b_{\mathbf{k}+\mathbf{G}_1} \\ b_{\mathbf{k}+\mathbf{G}_2} \\ b_{\mathbf{k}+\mathbf{G}_3} \\ \vdots \end{pmatrix}, \tag{5.50}$$

where $T(\mathbf{k}) = \hbar^2 k^2/(2m_e)$. The matrix in Eq. (5.50) must be diagonalized for each (fixed) $\mathbf{k}$ in the first BZ. Diagonalization is of course trivial whenever all off-diagonal matrix elements are identically null, i.e. for a constant potential $[\tilde{V}_{\text{eff}}(\mathbf{G} \neq \mathbf{0}) = 0]$: the eigenvalues in the solid are then simply the free-particle energies $\mathscr{E}_{\mathbf{k}}$ shifted by the constant potential $\tilde{V}_{\text{eff}}(\mathbf{0})$, and the eigenstates coincide with the original plane waves $|\mathbf{k}\rangle$. However, in any realistic solid, many $\mathbf{G} \neq \mathbf{0}$ Fourier components of V_{eff} are nonzero: these generate off-diagonal couplings among plane waves. The exact eigenstates of the problem are linear combinations of the plane waves differing by $\mathbf{G}$ vectors, obtained by the diagonalization of the full matrix in Eq. (5.50). As the $\mathbf{G}$-points are infinite, this is again an infinite matrix, but one can cut the basis restricting it to a finite number n_{pw} relevant plane waves separated by $\mathbf{G}$-points,

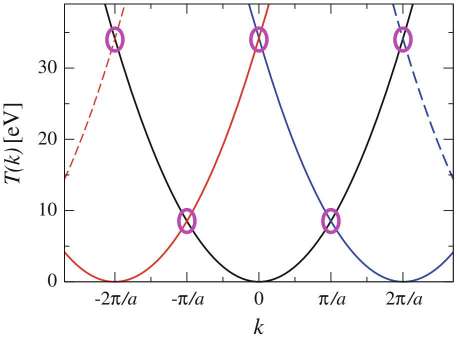


Fig. 5.49 Graphical intersections of a few translated free-electron parabolas $T(\mathbf{k}+\mathbf{G}_1)$, $T(\mathbf{k}+\mathbf{G}_2)$,... In a neighborhood of the *highlighted intersection points*, even small off-diagonal $\tilde{V}_{\text{eff}}$ ($\mathbf{G} \neq \mathbf{0}$) matrix elements can distort the parabolas quite significantly. In this figure, $a = 210\,\text{pm}$, and $G_l = l\,2\pi/a$

and then diagonalize numerically a finite version of Eq. (5.50). This method is used routinely for standard band-structure calculations.

Whenever the potential is "weak" (small $|\tilde{V}_{\text{eff}}(\mathbf{G} \neq 0)|$), analytic information can be extracted out of Eq. (5.50). More precisely, whenever the off-diagonal elements $\tilde{V}_{\text{eff}}(\mathbf{G}_2-\mathbf{G}_1)$ are small compared to the diagonal separation $|T(\mathbf{k}+\mathbf{G}_1)-T(\mathbf{k}+\mathbf{G}_2)|$ of the coupled states, the off-diagonal term acts as a small perturbation, thus it "perturbs" the diagonal energy only weakly, see Appendix B.9. As a result, if *all* couplings $\tilde{V}_{\text{eff}}(\mathbf{G}-\mathbf{G}_1)$ of a state $|\mathbf{k}+\mathbf{G}_1\rangle$ to *all* other plane-wave states are small compared to their diagonal energy separation, we can safely assume that the exact band energy shall not differ much from the diagonal energy

$$\mathcal{E}_{\mathbf{k}} \approx \frac{\hbar^2|\mathbf{k}|^2}{2m_e} + \tilde{V}_{\text{eff}}(\mathbf{0}). \tag{5.51}$$

However, even in the favorable case of small Fourier components of V_{eff}, the condition

$$\left|\tilde{V}_{\text{eff}}(\mathbf{G}_1-\mathbf{G}_2)\right| \ll |T(\mathbf{k}+\mathbf{G}_1) - T(\mathbf{k}+\mathbf{G}_2)| \tag{5.52}$$

is not verified for a set of special $\mathbf{k}$ points, namely those where two or more translated kinetic-energy parabolas get degenerate or nearly so: $T(\mathbf{k}+\mathbf{G}_1) \simeq T(\mathbf{k}+\mathbf{G}_2)$. At these points (highlighted in Fig. 5.49), the off-diagonal term $\tilde{V}_{\text{eff}}(\mathbf{G}_1-\mathbf{G}_2)$ becomes dominating, and it displaces the actual band significantly away from the free-electron parabola Eq. (5.51). At $\mathbf{k}$ points where states $|\mathbf{k}+\mathbf{G}_1\rangle$ and $|\mathbf{k}+\mathbf{G}_2\rangle$ are degenerate or nearly so, approximate band energies can be calculated by diagonalizing the 2×2 sub-matrix

$$\begin{pmatrix} T(\mathbf{k}+\mathbf{G}_1) + \tilde{V}_{\text{eff}}(\mathbf{0}) & \tilde{V}_{\text{eff}}(\mathbf{G}_1-\mathbf{G}_2) \\ \tilde{V}_{\text{eff}}(\mathbf{G}_2-\mathbf{G}_1) & T(\mathbf{k}+\mathbf{G}_2) + \tilde{V}_{\text{eff}}(\mathbf{0}) \end{pmatrix}.$$

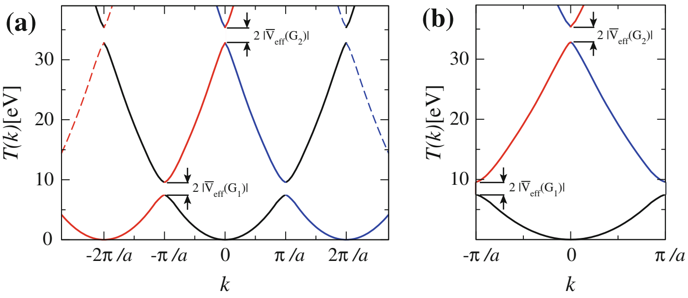


Fig. 5.50 **a** A gap of width $2|\tilde{V}_{\rm eff}(G)|$ opens at each degeneracy point $k = G_l/2$: the parabolic bands distort in an entire neighborhood around that k point. This representation adopts a repeated-zone scheme, which emphasizes that the bands are periodic functions of k, with the periodicity of the reciprocal lattice. **b** The same band structure, restricted to the first BZ

The spectrum of this matrix is discussed in Appendix B.5.2. The eigenenergies of Eq. (B.39), with $T(\mathbf{k}+\mathbf{G}_1)+\tilde{V}_{\rm eff}(\mathbf{0})$ in place of E_L, $T(\mathbf{k}+\mathbf{G}_2)+\tilde{V}_{\rm eff}(\mathbf{0})$ in place of E_R, and $|\tilde{V}_{\rm eff}(\mathbf{G}_1-\mathbf{G}_2)|$ in place of $-\Delta$, solve the 2×2 secular problem at hand. Figure B.2 shows that due to characteristic "repulsion" produced by the off-diagonal element, the band energies never come any closer than $\left|2\tilde{V}_{\rm eff}(\mathbf{G}_1-\mathbf{G}_2)\right|$.

This model illustrates the tendency of the periodic components of the potential to open forbidden energy intervals in the otherwise uninterrupted parabolic free-electron dispersion. As illustrated in Fig. 5.50, in 1D, gaps are guaranteed to open at all $\mathbf{G}/2$ points (unless some Fourier component $|\tilde{V}_{\rm eff}(\mathbf{G})|$ of the potential happens to vanish). In 3D, degeneracies of the kinetic term occur for all $\mathbf{k}$ such that $|\mathbf{k}+\mathbf{G}|=|\mathbf{k}|$, which is the condition (5.19) for Bragg scattering. The corresponding $|\tilde{V}_{\rm eff}(\mathbf{G})|$ opens a gap at a location in $\mathbf{k}$-space. In 3D, the periodic potential does not always generate a true band gap, i.e. a range of forbidden energy, since the forbidden energies for one $\mathbf{k}$-direction may well be allowed in some other $\mathbf{k}$-direction (see Fig. 5.54).

The free-electron starting point described here, like the tight-binding method of Sect. 5.2.1.1, lead to single-electron spectra characterized by bands of allowed energy separated by gaps of forbidden energy, in accord with Bloch's theorem. These models provide a physical meaning to the band index j of the electronic states in solids, see Eq. (5.26) and Fig. 5.42. In the tight-binding scheme, the band index contains indications mainly about the atomic labeling n of the state (which is a clear-cut concept especially for the narrow bands of inner shells, see Fig. 5.47). In the quasi-free-electron approach (especially fit for the wide bands related to the empty atomic levels), the band index j counts the number of times that the $\mathbf{k}\neq\mathbf{0}$ Fourier components of the potential have reflected the free-electron parabola back inside the first BZ.

5.2.2 *Filling of the Bands: Metals and Insulators*

The $T = 0$ (ground) state of a system of many independent electrons in a periodic potential is obtained by filling the one-electron band states up to a Fermi energy, like in the free-fermion model described in Sect. 4.3.2.1. The Fermi energy separates filled (below) from empty (above) levels. Depending on the total number of electrons in the solid, the Fermi energy may end up either inside one (several) energy band(s), or within a band gap (Fig. 5.51). In the first case, the electrons in the partly filled band(s) close to the Fermi energy are ready to take up excitation energy, provided typically by an external electromagnetic field. In particular, an arbitrarily weak applied electric field can accelerate electrons, which can then conduct electric current. Such a solid is a *metal*. In contrast, all electrons in completely filled bands are "frozen" by Pauli's principle. Any dynamical response of these electrons requires an excitation across some energy gap. A solid where all bands are either entirely filled or empty, with the Fermi energy inside a gap, is an *insulator*. In solids, the highest completely filled band is called *valence band*, while the lowest empty (or partly filled) band is the *conduction band*.

This basic difference affects substantially all properties of these two classes of materials, even at finite temperature. The Fermi-Dirac distribution (4.97) applies to independent electrons at equilibrium in a solid pretty much like in a free gas, simply replacing the free-electron energies and plane-wave states with the band energies and Bloch states. Thus, a nonzero temperature in a metal generates a finite concentration of electrons above the chemical potential and holes below it, similarly to what happens around the Fermi sphere of a free-electron gas. Accordingly, the thermodynamics of electrons in metals is interesting and rich of physical consequences, including a

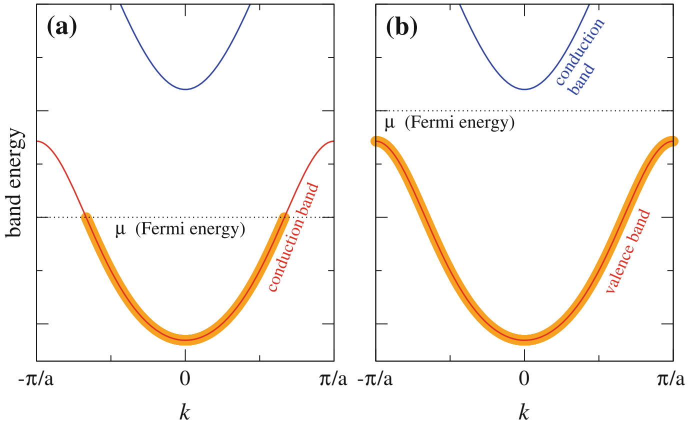
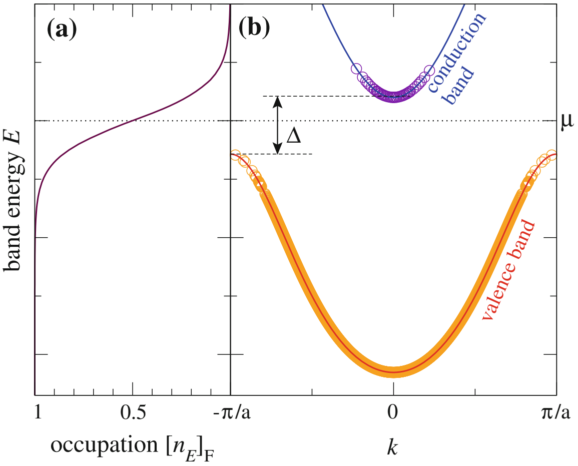


Fig. 5.51 The two basic $T = 0$ band-filling schemes: **a** a metal—the Fermi level (*dotted*) crosses a partly filled conduction band; **b** an insulator—the Fermi energy sits in a gap between a completely empty conduction band and a completely filled valence band

Fig. 5.52 A sketch of the occupation of the band levels (**b**) of a semiconductor, according to the finite-temperature Fermi distribution (**a**). The chemical potential μ sits in between the conduction-band bottom $\mathscr{E}_c$ and the valence-band top $\mathscr{E}_v$



characteristic T-linear contribution to the heat capacity of the solid, as discussed in Sect. 4.3.2.1. Instead, in an insulator of gap Δ between conduction and valence band, the average occupation of a valence-band state is $[n_v] \simeq 1 - \exp[-\Delta/(2k_B T)]$, extremely close to 1 at low temperature, and correspondingly the average occupation of a conduction-band state is $[n_c] \simeq \exp[-\Delta/(2k_B T)]$, extremely small at low temperature. For a temperature much smaller than Δ/k_B, the electrons of an insulator can be considered to all effects as frozen in the filled bands, their excitation being accessible only by means of high-energy spectroscopies. For a wide-gap insulator (e.g. Al_2O_3 $\Delta \simeq 5\,\text{eV}$, C diamond $\Delta \simeq 5.5\,\text{eV}$, SiO_2 $\Delta \simeq 8.0\,\text{eV}$, NaCl $\Delta \simeq 8.97\,\text{eV}$), any practical temperature is by far smaller than Δ/k_B, and the band occupancies are indistinguishable from those at $T = 0$. For example, with a 4 eV gap at room temperature ($k_B T \approx 0.025\,\text{eV}$), $[n_c] \approx e^{-80} \simeq 10^{-35}$. However, small-gap insulators, usually called *semiconductors* (e.g. Ge $\Delta \simeq 0.74\,\text{eV}$, Si $\Delta \simeq 1.17\,\text{eV}$, GaAs $\Delta \simeq 1.52\,\text{eV}$), show measurable conduction associated to thermal electronic excitations across the gap, even at room temperature, as sketched in Fig. 5.52, and discussed in Sect. 5.2.2.2.

But how can we predict the Fermi-energy position? In a many-electron atom, we just need to count electrons, but an infinite crystal boasts an infinite number of band states and of electrons: to determine the expected Fermi-energy position, we need a counting rule. Consider a macroscopic portion of the solid of volume $V = N_{n1} N_{n2} N_{n3} V_c$, extending for N_{ni} lattice repetitions in the $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ primitive directions. Apply periodic boundary conditions to this portion of solid, so that discrete translational invariance is preserved. To satisfy these artificial boundary conditions $\psi_{\mathbf{k}\,j}(\mathbf{r}) = \psi_{\mathbf{k}\,j}(\mathbf{r} + N_{ni}\mathbf{a}_i)$, the $\mathbf{k}$ label of the Bloch states is restricted to

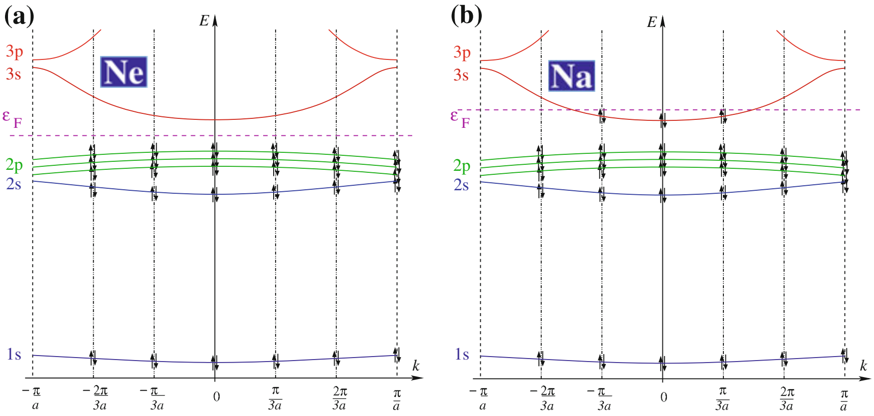


Fig. 5.53 The filling of the bands of a hypothetical Ne (**a**) and Na (**b**) 1D crystal composed of $N_n = 6$ atoms. According to Eq. (5.53), the allowed k-points in the first BZ are $k = 0, \pm\frac{1}{6}\frac{2\pi}{a}, \pm\frac{2}{6}\frac{2\pi}{a}$, and $\frac{3}{6}\frac{2\pi}{a}$. The $-\pi/a$ point must be excluded, since it coincides with π/a. Individual bands are assumed not to overlap. Energies are purely qualitative and not in scale

$$\mathbf{k} = \frac{n_1}{N_{n1}}\mathbf{b}_1 + \frac{n_2}{N_{n2}}\mathbf{b}_2 + \frac{n_3}{N_{n3}}\mathbf{b}_3,$$
$$\text{with } n_j = -\frac{N_{nj}}{2} + 1, -\frac{N_{nj}}{2} + 2, \ldots, \frac{N_{nj}}{2} - 2, \frac{N_{nj}}{2} - 1, \frac{N_{nj}}{2}. \qquad (5.53)$$

These values of $\mathbf{k}$ are the lattice equivalent to those of Eq. (4.38), and the 3D generalization of Eq. (5.27). These $N_n = N_{n1}N_{n2}N_{n3}$ discrete $\mathbf{k}$ values become dense and fill the primitive unit cell of the reciprocal lattice as $N_{nj} \to \infty$ and the infinite real-space crystal is recovered (see Fig. 5.44). The number of electrons of this finite crystal portion equals N_n times the number n_{cell} of electrons of each unit cell. Each band (orbital) state makes room for 2 electrons, one for each spin state, ↑ and ↓. If the bands are all disjoint, one on top of another, the $N_n n_{\text{cell}}$ electrons fill the $2N_n$ spin-orbital states of the $n_{\text{cell}}/2$ lowest bands, as illustrated in Fig. 5.53. An even value of n_{cell} leads to $n_{\text{cell}}/2$ full bands, followed by empty bands above. In the example of Fig. 5.53, each atom of solid neon (fcc, one atom per cell) carries $n_{\text{cell}} = 10$ electrons to the bands, for a total of $10N_n$ electrons in the crystal. In a tight-binding language, $2N_n$ electrons fill the 1s band, $2N_n$ electrons fill the 2s band, and $6N_n$ electrons fill completely the 2p bands, for a total of 5 filled bands. The Fermi energy then lies in the gap between the filled 2p band and the 3s-3p empty bands above: the neon crystal is an insulator.

Odd n_{cell} leads to $(n_{\text{cell}} - 1)/2$ full bands plus one half-filled band. For example, the $n_{\text{cell}} = 11$ electrons that each sodium atom puts in the band states of its bcc crystal (one atom per cell) fill completely the 1s, 2s, 2p bands, and fill only the $N_n/2$ lowest-energy spin-orbital states of the 3s band. This 3s band is then half filled, the Fermi energy cutting through it: the sodium solid is thus a metal.

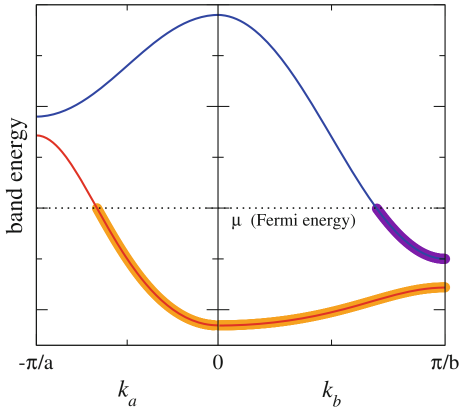


Fig. 5.54 Band overlap. The band energies $\mathcal{E}_{\mathbf{k}\,j}$ of 2D and 3D crystals depends on the **k** direction. Even with a gap in the band structure in each **k**-direction, when taking all directions into account the overall gap can vanish, since the energy range forbidden in some direction k_a becomes allowed in a different direction k_b: these bands overlap. In this kind of solid, even with an even number of electrons per cell, the Fermi energy can thus end up crossing several bands: that solid is a metal

While it is evident that *any band crystal with an odd* number of electrons per cell n_{cell} *is a metal*,[16] for even n_{cell} both insulators (like Ne) and metals are possible, since the $(n_{\text{cell}}/2)$-th and the $(n_{\text{cell}}/2 + 1)$-th bands are not necessarily separated by a gap, as illustrated in Fig. 5.54. The alkali earth (IIA) and end of the transition (IIB: Zn, Cd, and Hg) elemental solids are all metals, with even n_{cell}, precisely due to overlapping bands at the Fermi energy.

5.2.2.1 Metals

The distinctive feature of a metal is its ability to conduct electric current. In practice all solids, even insulators, show some measurable conductivity associated to impurities and thermal excitations. What does characterize uniquely the metallic state is a conductivity which decreases as T is increased, contrasted to the conductivity of insulators, which increases as T is increased due to extra thermally excited charge carriers. In the **k** space of metals a *Fermi surface* separates full and empty states at $T = 0$, like the Fermi sphere of the ideal Fermi gas.

We now analyze the predictions of band theory for the movement of electrons in crystals. To describe the motion of electrons in the periodic field of the ions plus the applied external field, a semiclassic approach is useful. The standard approach of QM is to represent an electron as a wave packet, i.e. a superposition of Bloch states of a single band j, characterized by a peaked wave-number distribution, thus a large spatial extension (much larger than the crystal lattice spacing, see Fig. 5.55).

---

[16] When the independent electrons approximation breaks down, as in *strongly correlated materials*, there occur insulating states with odd n_{cell}, often accompanied by magnetic order of the spins.

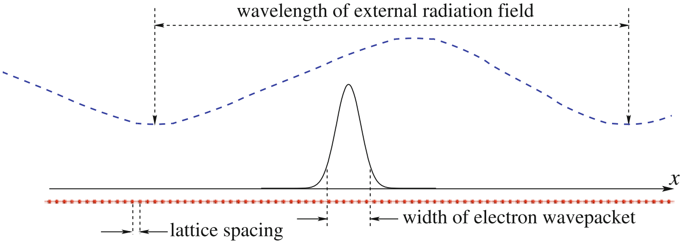


Fig. 5.55 A cartoon for the electron dynamics within the semiclassic model. The length over which externally applied fields (*dashed*) vary far exceeds the spread in the electron wavepacket (*solid*), and this wavepacket, in turn, extends over a region several lattice constants wide

The following equations govern the motion of the center of mass $\mathbf{r}$ of such a wave packet [10]:

$$\frac{d}{dt}\mathbf{r} = \mathbf{v}_j(\mathbf{k}) = \frac{1}{\hbar}\nabla_{\mathbf{k}}\,\mathscr{E}_{\mathbf{k}\,j} \tag{5.54}$$

$$\hbar\frac{d}{dt}\mathbf{k} = -q_e\left[\mathbf{E}(\mathbf{r},t) + \mathbf{v}_j(\mathbf{k}) \times \mathbf{B}_{\text{ext}}(\mathbf{r},t)\right]. \tag{5.55}$$

The external perturbing electric $\mathbf{E}(\mathbf{r}, t)$ and magnetic $\mathbf{B}_{\text{ext}}(\mathbf{r}, t)$ fields[17] are supposed to vary slowly on the scale of the wave-packet size (Fig. 5.55). A rigorous derivation of Eqs. (5.54) and (5.55) goes beyond the scope of the present course: we suggest a few heuristic arguments to support their plausibility.

- The center-mass velocity $\mathbf{v}_j$ of the wave packet is the *group velocity* associated to the dispersion $\mathscr{E}_{\mathbf{k}\,j}$ of the Bloch waves. Equation (5.54) states the basic fact of wave mechanics that a wave packet with dispersion $\omega(\mathbf{k}) = \hbar^{-1}\mathscr{E}_{\mathbf{k}\,j}$ moves with the velocity given by its group velocity $\nabla_{\mathbf{k}}\,\omega(\mathbf{k})$. In the special case of a free electron $\mathscr{E}_{\mathbf{k}} = \hbar^2 k^2/(2m_e)$, this yields the usual relation $\mathbf{v}(\mathbf{k}) = \hbar\mathbf{k}/m_e$ between velocity and momentum.
- If the force associated to a static external electric potential $\phi_{\text{ext}}(\mathbf{r})$ acts on the band electron, then its total energy $\left[\mathscr{E}_{\mathbf{k}\,j} - q_e\,\phi_{\text{ext}}(\mathbf{r})\right]$ should remain conserved along the semiclassic motion. To verify this, we derive this total energy with respect to time (notation: $\frac{df}{dt} \equiv \dot{f}$):

$$\frac{d}{dt}\left[\mathscr{E}_{\mathbf{k}\,j} - q_e\,\phi_{\text{ext}}(\mathbf{r})\right] = \nabla_{\mathbf{k}}\,\mathscr{E}_{\mathbf{k}\,j}\cdot\dot{\mathbf{k}} - q_e\,\nabla_{\mathbf{r}}\,\phi_{\text{ext}}(\mathbf{r})\cdot\dot{\mathbf{r}} = \mathbf{v}_j(\mathbf{k})\cdot\left[\hbar\dot{\mathbf{k}} + q_e\,\mathbf{E}(\mathbf{r})\right],$$

---

[17] The magnetic field $\mathbf{H}$ produced by currents *external* to the solid is related to the magnetic induction field $\mathbf{B}_{\text{ext}}$ of Eq. (5.55) by $\mathbf{H} = \varepsilon_0 c^2\,\mathbf{B}_{\text{ext}}$.

where we use $\nabla_{\mathbf{k}}\,\mathcal{E}_{\mathbf{k}\,j} = \hbar\mathbf{v}_j(\mathbf{k})$, Eq. (5.54). If Eq. (5.55) is satisfied, then this derivative indeed vanishes, as the vector $\hbar\dot{\mathbf{k}}+q_e\mathbf{E} = -q_e\mathbf{v}_j\times\mathbf{B}_{\text{ext}}$ is perpendicular to the velocity $\mathbf{v}_j$. Total energy is thus conserved.

- Equation (5.55) recalls the classical equation of motion of a particle of charge $-q_e$ moving under the action of the *external* electromagnetic fields $\mathbf{E}(\mathbf{r})$ and $\mathbf{B}_{\text{ext}}(\mathbf{r})$ only. The periodic forces produced by the crystal act through the band dispersion $\mathcal{E}_{\mathbf{k}\,j}$, generating the nontrivial $\mathbf{k}$-dependence of the velocity in Eq. (5.54), replacing the free-electron $\mathbf{v} = \hbar\mathbf{k}/m_e$. This means that in the crystal $\hbar\mathbf{k}$ *does not equal the electron momentum*, as for a free electron. It is rather called *crystal momentum*.
- The semiclassic equations assume that the external fields are sufficiently weak to induce no inter-band transitions. Strong fields would make the semiclassic approximation fail and lead to electric or magnetic *breakdown*. Also, the single-band semiclassic equations hold until the frequency of any time-oscillating field does not approach inter-band gaps ($\omega \ll \varDelta/\hbar$). Rapidly varying fields are applied in spectroscopy precisely with the purpose of inducing inter-band transitions.

According to the semiclassic equations, the electrons of a completely filled band do not contribute to either electric or heat current. The electric *current density* carried by a wave packet representing an electron moving in a volume V is $(-q_e)\mathbf{v}_j(\mathbf{k})/V$. The total electric current density carried by all electrons in a filled band j amounts to[18]

$$\mathbf{j} = \int_{\text{BZ}} (-q_e)\mathbf{v}_j(\mathbf{k})\,\frac{1}{4\pi^3}\,d^3\mathbf{k} = \mathbf{0}. \tag{5.56}$$

This integral vanishes since, due to Eq. (5.54), the integrand function is the gradient of a periodic function ($\nabla_{\mathbf{k}}\mathcal{E}_{\mathbf{k}\,j}$, and $\mathcal{E}_{\mathbf{k}\,j}$ is periodic with the reciprocal-lattice periodicity), which is integrated over an *entire cell* (e.g. the first BZ). In other terms, in a filled band for each electron carrying current in some direction another electron carries current the opposite way, totaling a vanishing net current. The same observations apply to the energy (heat) current

$$\mathbf{j}_{\mathcal{E}} = \int_{\text{BZ}} \mathcal{E}_{\mathbf{k}\,j}\mathbf{v}_j(\mathbf{k})\,\frac{1}{4\pi^3}\,d^3\mathbf{k} = \mathbf{0}, \tag{5.57}$$

by noting that the integrand is proportional to $\nabla_{\mathbf{k}}(\mathcal{E}_{\mathbf{k}\,j})^2$. Completely filled band do not contribute to transport any more than completely empty bands. All *electric and thermal conductivity is to be attributed to partly filled bands*. This explains why no systematic increase of conductivity is observed in the crystalline elements for increasing Z (for example the conductivities of fcc Cu, Ag, and Au are quite similar), despite the largely different total number of electrons.

---

[18] Similarly to Eq. (4.90), the n_i sum is turned into an integral by inserting the appropriate density of states. According to Eq. (4.38), the $\mathbf{k}$-density of states is $V/(2\pi)^3$. An extra factor $g_s = 2$ accounts for the spin degeneracy. The integration of $V/(4\pi^3)$ over the whole BZ represents the summation over all states in band j.

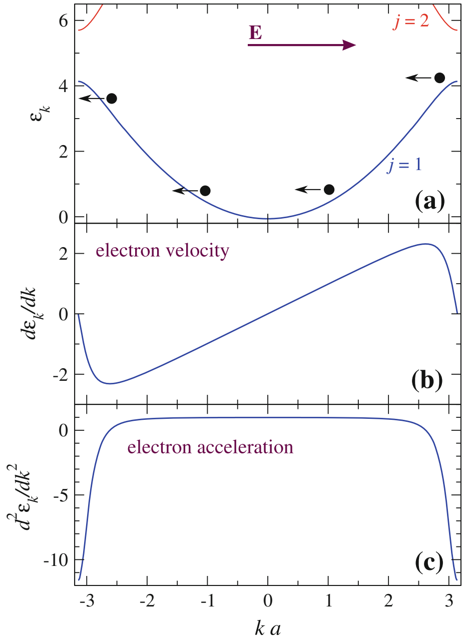


Fig. 5.56 The reciprocal-space motion of the electrons of a 1D crystal. Under the action of a leftward external force (uniform constant rightward electric field **E**), the wave number k drifts at constant speed. At the BZ boundary k can be brought back into the first BZ by $k = -\pi/a \to \pi/a$, and then continues to move leftward and traverse the BZ again and again. Correspondingly, **a** the electron band energy oscillates. The real-space electron **b** velocity and **c** acceleration also oscillate, because they are proportional to the first, Eq. (5.54), and second, Eq. (5.64), derivative of the band energy

On the other hand, Bloch states are *stationary* states of the Schrödinger equation in the perfect crystal: if a wave packet of Bloch states representing an electron has a finite mean velocity (as happens unless by chance $\nabla_{\mathbf{k}} \mathscr{E}_{\mathbf{k}\,j} = \mathbf{0}$), then that velocity shall persist forever. Thus, this semiclassic theory for Bloch states predicts that, even in the absence of any external electric field, metals should carry persistent currents. No such persistent currents are observed. Even worst, as illustrated in Fig. 5.56 for the simple 1D case, the semiclassic motion following Eq. (5.55) under the action of a constant field cycles the k-point across the whole BZ. Correspondingly, Eq. (5.54) yields an oscillating velocity generating positive and negative currents for the same amount of time. A DC electric field should then induce an AC current (*Bloch oscillations*) in a metal wire! This unphysical prediction is neither an artifact of 1D nor of the semiclassic approximation: real metals under ordinary conditions exhibit a completely different response: *Ohm's law*[19] $\mathbf{j} = \sigma \mathbf{E}$. This inconsistency of the model

---

[19] The current I in a wire is directly proportional to the applied potential drop: $I = R^{-1} V$. The coefficient of proportionality R depends on the length L and cross section A of the wire, but not on the current or potential drop. In terms of the current density **j** crossing perpendicularly the surface

with observation is due to Bloch electrons (and wave packets thereof) being capable to travel forever through a perfect crystal, without any energy dissipation. Any real crystal deviates from ideality because of

- structural defects, as discussed in Sect. 5.1;
- its nuclei not being frozen at their equilibrium positions but actually vibrating around them, as we shall discuss in Sect. 5.3.

Both these discrepancies from the ideal-crystal picture are sources of *collisions* for conduction electrons.[20] To represent the effect of collisions as simply as possible, we shall assume that:

1. Each electron experiences an instantaneous collision at random, with probability τ^{-1} per unit time. The time τ, variously known as the *relaxation time*, the *collision time* or the *mean free time*, represents the average time that an electron travels freely between a collision and the next.
2. The electron emerges from a collision in a random $\mathbf{k}$ state, reflecting the (Fermi) distribution at the appropriate local temperature, and respecting Pauli's principle. All memory of the initial $\mathbf{k}$ prior to collision is lost.
3. Between successive collisions, each electron moves according to the semiclassic equations of motion (5.54), (5.55).

As a result of frequent collisions, the external field induces only a weak perturbation to the thermal equilibrium distribution. Basically, collisions act mostly close to the *Fermi surface*. When, following Eq. (5.55), a DC electric field attempts to shift each occupied $\mathbf{k}$ state in the $-q_e\mathbf{E}$ direction, collisions rapidly transfer electrons from occupied states in the higher-energy region back into the emptied region of lower energy. This tendency of collisions to re-establish thermal equilibrium, illustrated in Fig. 5.57a, quenches the free acceleration of the electrons, thus preventing the Bloch oscillations. After an initial transient, the net effect of the field $\mathbf{E}$ amounts to a small steady displacement of the filled states relative to zero field (Fig. 5.57b). This displacement generates a steady current density $\mathbf{j}$ equal to the $\mathbf{k}$-space integration of $V^{-1}\,(-q_e)\mathbf{v}_j(\mathbf{k})$ through the region $\delta^3 k$ of unbalanced occupation. $\mathbf{j}$ can be estimated roughly for a free-electron parabolic band as

$$\mathbf{j} = \int_{\delta^3 k} V^{-1}\,(-q_e)\mathbf{v}_j(\mathbf{k})\,\frac{V}{4\pi^3} d^3\mathbf{k} \simeq \int_{\delta^3 k} (-q_e)\hat{\mathbf{E}} v_{\mathrm{F}}\, d^3\mathbf{k} \simeq -q_e\, v_{\mathrm{F}}\,\hat{\mathbf{E}}\,(\delta^3 k), \quad (5.58)$$

(Footnote 19 continued)
area A, $I = |\mathbf{j}|A$. In terms of the electric field $\mathbf{E}$, the potential drop $V = L|\mathbf{E}|$. Thus Ohm's law rewrites $\mathbf{j} = L/(AR)\,\mathbf{E} = \sigma\,\mathbf{E}$, where the conductivity $\sigma = L/(AR)$ and resistivity $\rho = \sigma^{-1}$ are characteristic properties of the material,.

[20] Some scattering is produced by electron-electron interaction as well, but this mechanism is quite negligible compared to the two other sources of collisions in ordinary metals at ordinary temperature.

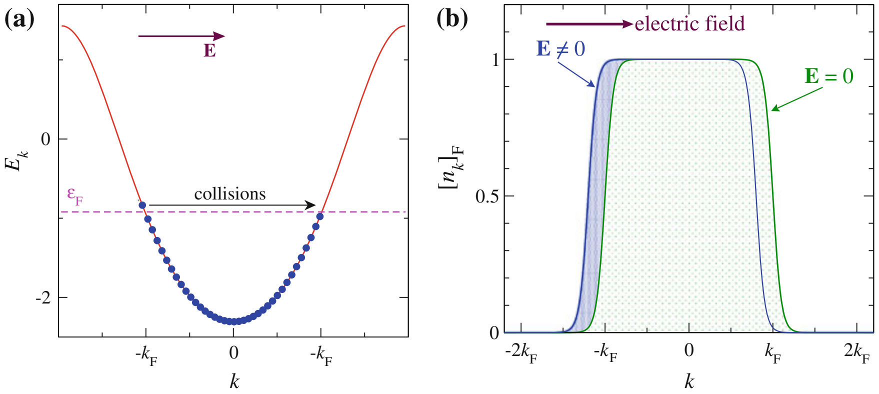


Fig. 5.57 **a** Band occupation under the combined effect of (i) a rightward static external electric field **E**, accelerating the electrons to the left according to Eq. (5.55), and (ii) collisions with crystal defects and vibrations, which tend to reestablish equilibrium by scattering extra-energetic electrons prevalently into lower-energy states which have been left empty. **b** Under the combined effect of the field and collisions, the occupation distribution shifts to the left in **k** space. This steady shift, here largely exaggerated, is proportional to both |**E**| and the average time τ between collision, and accounts for electric-current transport

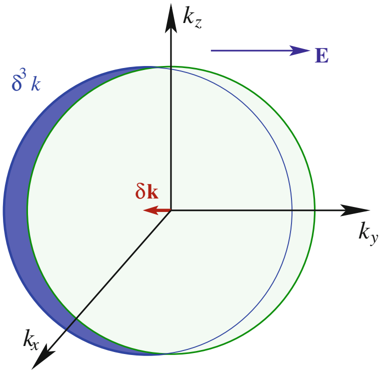


Fig. 5.58 The shift $\delta\mathbf{k}$ of the Fermi sphere induced by a rightward external electric field $\mathbf{E} = E\,\hat{\mathbf{y}}$. The highlighted **k**-space volume $(\delta^3 k)$ is responsible for the velocity-distribution asymmetry supporting the net electric current carried by this partly filled band

where we drop numerical factors of order 1. The **k**-space volume $(\delta^3 k)$ in between the equilibrium Fermi surface and its field-shifted replica, sketched in Fig. 5.58 for free electrons, can be estimated by observing that in the average time τ between two collisions, each electron changes its wave vector by $\delta\mathbf{k} = \hbar^{-1}\tau\,(-q_e)\,\mathbf{E}$. Dropping again factors of order unity, the volume $(\delta^3 k)$ is approximately

$$(\delta^3 k) \simeq -|\delta\mathbf{k}|\, k_F^2 \simeq \frac{\tau}{\hbar}\,(-q_e)\,|\mathbf{E}|\; k_F^2,$$

where the minus sign indicates that the shift is opposite to the **E** field. By substituting this expression and $v_F \simeq \hbar k_F/m_e$ (valid for free electrons) in Eq. (5.58), we obtain

Table 5.2 The measured electrical resistivity $\rho = \sigma^{-1}$ of a few elemental metals [10]

Element	ρ (nΩm) at 77 K	ρ (nΩm) at 273 K	τ (10^{-14} s) at 77 K	τ (10^{-14} s) at 273 K
Na	8	42	17	3.2
K	14	61	18	4.1
Rb	22	110	14	2.8
Cu	2	16	21	2.7
Ag	3	15	20	4.0
Au	5	20	12	3.0
Mg	6	39	7	1.1
Al	3	25	6.5	0.8

The corresponding relaxation times are obtained from Eq. (5.60) through $\tau = m_e / \left(\rho q_e^2 \frac{N}{V}\right)$, with $\frac{N}{V}$ equaling the density of conduction electrons (Data from Ref. [10])

$$\mathbf{j} \simeq (-q_e)\frac{\hbar k_{\rm F}}{m_e}\,\hat{\mathbf{E}}\,\frac{\tau}{\hbar}\,(-q_e)\,|\mathbf{E}|\,k_{\rm F}^2 \simeq \frac{q_e^2\,\tau}{m_e}\,k_{\rm F}^3\,\mathbf{E} \simeq \frac{q_e^2\,\tau}{m_e}\,\frac{N}{V}\,\mathbf{E}, \tag{5.59}$$

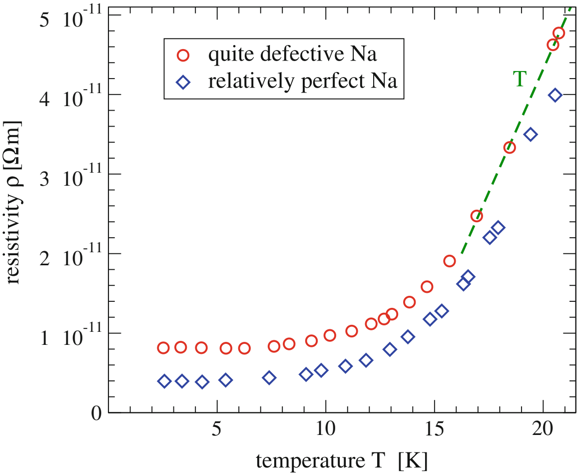
where we used the free-electron relation of $k_{\rm F}$ with electron density, $k_{\rm F}^3 = 3\pi^2 N/V$ (Sect. 4.3.2.1), dropping factors of order unity. Equation (5.59) agrees with Ohm's law, and yields a conductivity

$$\sigma \simeq \frac{q_e^2\,\tau}{m_e}\,\frac{N}{V}. \tag{5.60}$$

By measuring the resistivity $\rho = \sigma^{-1}$ and the conduction-electron number density N/V of metals, we evaluate the relaxation time τ through Eq. (5.60). The values reported in Table 5.2, in the 10^{-13} s range, indicate that between two successive collisions an electron travels an average distance (called the electron's *mean free path*) $d = \tau v_{\rm F} \simeq 100$ nm, hundreds of typical interatomic spacings.

The observed trend of increasing resistivity with temperature, indicates that the relaxation time decreases. Collisions become more frequent as thermal motion produces larger displacements of the nuclei from their equilibrium positions, as described in Sect. 5.3. The average time between collisions is indeed inversely proportional to the total number of phonons $\tau^{-1} \propto \sum_\varepsilon [n_\varepsilon]_B$. At thermal energies $k_{\rm B}T$ much larger than the characteristic phonon energies ($\varepsilon \lesssim 100$ meV), the total phonon number is proportional to T (verify this by expanding Eq. (4.96): $[n_\varepsilon]_B = (e^{\beta\varepsilon} - 1)^{-1} \simeq (\beta\varepsilon)^{-1} = k_{\rm B}T/\varepsilon$). Indeed, at high temperature metals do exhibit a resistivity approximately $\rho \propto T$. At low temperature, the phonon number decreases rapidly [see Eq. (4.116)], but all sorts of crystal defects provide a residual T-independent scattering. Accordingly, at low temperature the T-linear regime turns into a T-independent (but sample-dependent) resistivity, as illustrated in Fig. 5.59

Fig. 5.59 The low-temperature resistivity of sodium, measured for two samples characterized by different defect concentrations, leading to different low-T residual resistivity. At low temperature the tiny phonon contribution to resistivity grows as T^5, but it then turns rapidly into a T-linear increase (*dashed line*)



for Na. The low-temperature resistivity contribution of defects varies widely among different metals, and is huge in disordered alloys.

In analogy to electrical conductivity, we can estimate the *thermal conductivity* of electrons in a metal. When a thermal gradient $\nabla_{\mathbf{r}} T = \hat{\mathbf{z}}\, {}^{dT}\!/_{dz}$ is established across the sample, electrons reaching a given point inside the metal, have a slightly different energy distribution, depending on whether they are coming from the high-temperature side or from the low-temperature side. This difference is seen mainly as a different broadening near the Fermi energy, and its amount depends on the average distance $d \simeq \tau v_{\mathrm{F}}$ that these electrons have traveled since the previous collision. As sketched in Fig. 5.60, the electrons coming to that point from a region at temperature $T + \Delta T$ carry an extra energy $V^{-1} C_v \,\Delta T = -V^{-1} C_v\, d\, \hat{\mathbf{z}} \cdot \nabla_{\mathbf{r}} T$ where C_v is the heat capacity, Eq. (4.106), and the minus sign indicates that energy is transported in the direction opposite to the temperature gradient. The heat current density is obtained by multiplying this heat density by the speed v_{F} at which electrons travel:

$$\mathbf{j}_{\mathscr{E}} \simeq -V^{-1} N k_{\mathrm{B}} \frac{\pi^2}{2} \frac{k_{\mathrm{B}} T}{\varepsilon_{\mathrm{F}}} \tau v_{\mathrm{F}} \nabla_{\mathbf{r}} T v_{\mathrm{F}} \simeq -k_{\mathrm{B}}^2\, T\, \tau \frac{v_{\mathrm{F}}^2}{\varepsilon_{\mathrm{F}}} \frac{N}{V} \nabla_{\mathbf{r}} T \simeq -\frac{k_{\mathrm{B}}^2\, T\, \tau}{m_e} \frac{N}{V} \nabla_{\mathbf{r}} T, \tag{5.61}$$

where we use the relation $\varepsilon_{\mathrm{F}} = m_e v_{\mathrm{F}}^2/2$, and we drop factors of order unity. The expression (5.61) for the heat current is compatible with a *thermal conductivity*[21]

$$\mathscr{K} \simeq \frac{k_{\mathrm{B}}^2\, T\, \tau}{m_e} \frac{N}{V}. \tag{5.62}$$

By comparing σ and $\mathscr{K}$ in Eqs. (5.60) and (5.62) we predict a ratio

---

[21] The underlying linear-response relation is $\mathbf{j}_{\mathscr{E}} = -\mathscr{K}\, \nabla_{\mathbf{r}} T$.

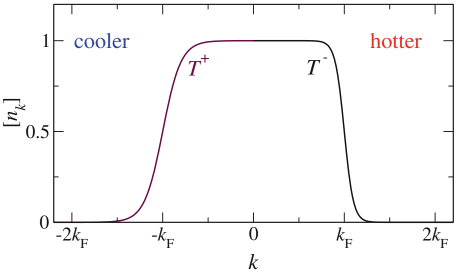


Fig. 5.60 Heat transport is associated to electrons moving out from the hotter (*right*) region carrying (on average) higher energy than those coming from the cooler (*left*) region. At a given point $\mathbf{r}$ in the metal, the distribution of $\mathbf{k}$ states has therefore a slight asymmetry (strongly exaggerated in this cartoon) in the direction of $\nabla_{\mathbf{r}} T$

$$\frac{\mathcal{K}}{\sigma} = \frac{\pi^2}{3} \frac{k_B^2}{q_e^2} T \tag{5.63}$$

between thermal and electric conductivities according to the relaxation-time model considered. A careful analysis of the factors of order unity previously ignored yields the factor $\pi^2/3$ in Eq. (5.62). This relation predicts that good electrical conductors are also good heat conductors. The empirical observation of this fact is known as the *Wiedmann-Franz law*. Experimentally, for a broad range of metals and temperature, the ratio $\mathcal{K}/\sigma$ follows Eq. (5.63) surprisingly well, given the model simplifications. Indeed, measurements of the ratio $\mathcal{K}/(\sigma T)$ normalized to the universal expression $\pi^2 k_B^2/(3q_e^2)$, yield: 0.868 (Na, 273 K), 0.950 (Au, 273 K), 0.966 (Au, 373 K), 1.08 (Pb, 273 K), 1.04 (Pb, 373 K), indicating moderate deviations from Eq. (5.63).

Other transport experiments, both in the DC and AC regime can be interpreted in terms of this simple relaxation-time model. An important class of measurements is related to the Hall effect: an electric field arising perpendicular to a current running through a sample immersed in a magnetic field $\mathbf{B}_{\text{ext}}$ also perpendicular to the current. The Hall field is generated by the charge accumulating on the conductor sides due the Lorentz force of Eq. (5.55), as sketched in Fig. 5.61. The Lorentz force acting on the carriers is independent of the sign of their charge. Therefore, the sign of the charge buildup and of the ensuing transverse Hall field probes the sign of the charge carriers. In most metals, the resulting Hall field is, as in Fig. 5.61, consistent with negative charge carriers. Surprisingly however, a number of metals (e.g. beryllium and cadmium) exhibit a reversed Hall field, as if their carriers had a positive charge! This inversion is one of the most spectacular consequences of the deviations of the crystalline bands from a simple free-electron parabola.

The main effect of the crystal periodic potential is to replace the free-electron dispersion (a parabola whose curvature $\hbar^2/m_e$ is fixed by the electron mass) with nontrivial bands. As illustrated in Fig. 5.56, electrons in a crystal band accelerate at a different rate compared to free electrons, due to nontrivial local band curvature. According to Eqs. (5.54) and (5.55),

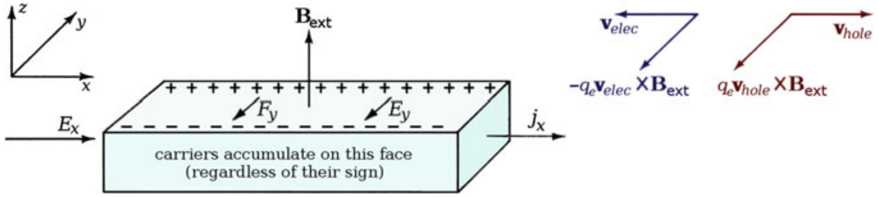


Fig. 5.61 A scheme of Hall's experiment: the Lorentz force deflects the charge carriers moving in a conductor immersed in a perpendicular magnetic field. This deflecting force F_y pushes carriers toward the sample front surface regardless of the sign of the charge carriers. In the steady state, the charge accumulated at the surface generates a transverse electric field E_y that balances the Lorentz force, and whose sign reflects the sign of the charge carriers: the scheme depicts the sign appropriate for electrons (Inspired by Fig. 1.3 of Ref. [10])

$$\begin{aligned}\frac{d^2}{dt^2}\mathbf{r} &= \frac{d}{dt}\mathbf{v}_j(\mathbf{k})\\ &= \left[\nabla_{\mathbf{k}}\mathbf{v}_j(\mathbf{k})\right]\cdot\frac{d\mathbf{k}}{dt} = \frac{1}{\hbar^2}\sum_{uw}\hat{\mathbf{e}}_u\left[\frac{\partial^2\mathscr{E}_{\mathbf{k}\,j}}{\partial k_u\partial k_w}\right]\frac{d(\hbar k_w)}{dt} = \sum_{uw}\hat{\mathbf{e}}_u(m^*)^{-1}_{uw}\,F_w,\end{aligned} \tag{5.64}$$

where $\hat{\mathbf{e}}_1 = \hat{\mathbf{x}}$, $\hat{\mathbf{e}}_2 = \hat{\mathbf{y}}$, $\hat{\mathbf{e}}_3 = \hat{\mathbf{z}}$ indicate the Cartesian versors, and $\mathbf{F} = -q_e(\mathbf{E} + \mathbf{v}_j \times \mathbf{B}_{\text{ext}})$ represents the total external force acting on the electron. The final form of Eq. (5.64) expresses a sort of Newton equation, with an inverse mass *tensor* of components

$$(m^*)^{-1}_{uv} = \frac{1}{\hbar^2}\frac{\partial^2\mathscr{E}_{\mathbf{k}\,j}}{\partial k_u\partial k_v}, \tag{5.65}$$

describing the band curvature. As the electrons active in transport are those near the Fermi energy, this curvature is to be evaluated at some $\mathbf{k}$ point at the Fermi surface. $|m^*|$ replaces the free-electron mass in Eqs. (5.60) and (5.62), thus accounting for the actual value of the acceleration of electrons close to the Fermi energy in a crystal-potential band. In 1D m^* is proportional to the inverse of the band curvature, Fig. 5.56c. In 3D, the effective mass m^* is a suitable average over the tensor components $(m^*)_{uv}$, and may differ substantially from the free-electron mass m_e. In particular m^* turns out larger than m_e for narrow flat bands, characterized by a weak curvature, such as 3d bands of transition metals or 4f bands of rare-earth metals. According to Eqs. (5.60) and (5.62), a larger mass leads to smaller conductivity because, in between two collisions, external fields can accelerate heavier band electrons less than free electrons. Note however that the carrier mass drops out in the ratio $\mathscr{K}/\sigma$ of Eq. (5.63): the Wiedmann-Franz law should and does hold roughly independently of m^*.

Importantly, *negative* effective masses occur whenever the Fermi level sits in a region where the band curves downward (e.g. close to the BZ boundary of Fig. 5.57a). An electron of negative m^* accelerates in a direction opposite to the applied force: as its charge $(-q_e)$ is negative, it accelerates in the same direction as the external electric field. A negative-m^* electron behaves therefore in all ways as a particle of positive charge $(+q_e)$ and positive mass $|m^*|$, called a *hole*. Holes carry electric current in the same direction as the applied field (like genuine electrons), but produce reversed Hall effect, because they behave as positive carriers. Thus, a negative effective mass, i.e. hole conduction, explains the reversed Hall field of several metals.

5.2.2.2 Semiconductors

Semiconductors are insulators characterized by a ≈1 eV gap between the valence and conduction bands. For example, Fig. 5.62 sketches the bands of solid Si and Ge, semiconductors sharing the same crystal structure as C diamond but with larger lattice spacing a. In these elemental semiconductors, the Fermi energy sits inside the bandgap between a full sp^3 bonding-type band and an empty sp^3 antibonding-type band. Note that the possibility of this band "splitting" is directly connected to the diamond crystal structure, and would not occur in a hypothetical simple-cubic or fcc

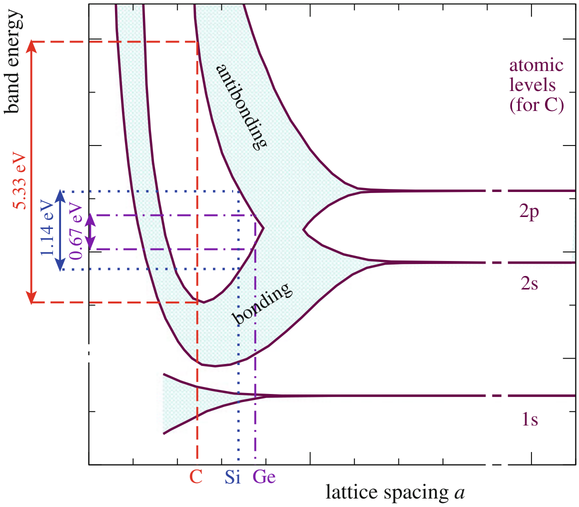


Fig. 5.62 Qualitative lattice-parameter dependence of the band energies for the IVB group solids, those with the diamond structure. The narrow separate s and p bands at unphysically large a expand and evolve into a hybrid sp^3 band as a is reduced. Eventually this hybrid band splits into a filled "bonding" and an empty "antibonding" band (similar to the bonding and antibonding states of CH$_4$). These bands move apart as the crystal shrinks ($a_{\mathrm{Ge}} > a_{\mathrm{Si}} > a_{\mathrm{C}}$). This electronic structure leads to the bandgap ordering $\Delta_{\mathrm{Ge}} < \Delta_{\mathrm{Si}} < \Delta_{\mathrm{C}}$, and to the uncommon bandgap increase under applied external pressure (and decrease with thermal expansion)

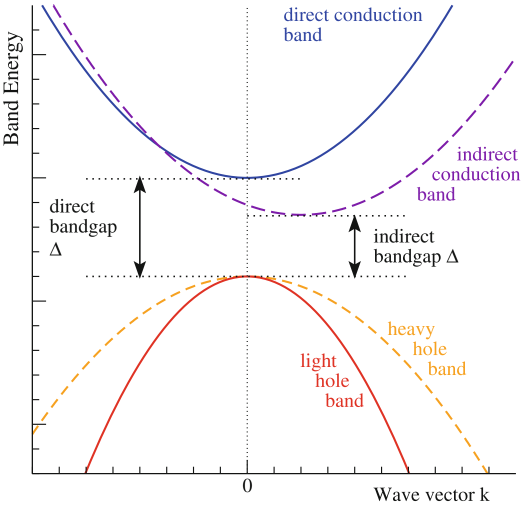


Fig. 5.63 Direct versus indirect gap in insulators and semiconductors. A direct gap can be probed spectroscopically by "direct" optical absorption. In contrast, absorption through an indirect gap instead must be assisted by some phonon absorption/emission in order to grant the wave-number conservation in the process

Si/Ge; precisely this bandgap yields a large cohesive stability to the rarefied diamond structure of those solids where the number of electrons matches the capacity of the bonding-type band.[22] Semiconductors of type III–V (e.g. GaAs) and some II–VI (e.g. BeSe) crystallize in a similar crystal structure, namely the zincblend structure, with two different chemical species occupying the two geometrically inequivalent sites of the unit cell of the diamond structure (Fig. 5.25). Different structures are observed in other semiconductors. Due to structural and chemical differences, the bands of individual compounds are qualitatively and quantitatively different. In particular, the gap $\Delta = \mathscr{E}_c - \mathscr{E}_v$ between the top of the valence band and the bottom of the conduction band can be either direct (same **k** for the minimum $\mathscr{E}_c$ of $\mathscr{E}_{\mathbf{k}\,c}$ and for the maximum $\mathscr{E}_v$ of $\mathscr{E}_{\mathbf{k}\,v}$, as in GaAs and InP) or indirect (when these two extrema occur at different **k** points, as in Si, Ge, GaP), see Fig. 5.63.

Transport in a pure (*intrinsic*) semiconductor is dominated by the thermal excitation of electrons from the valence into the conduction band. The average number density of such electrons is

$$N_c = \frac{[n_c]}{V} = \frac{1}{V}\int_{\mathscr{E}_c}^{\infty} g(\mathscr{E})\,[n_{\mathscr{E}}]_F\, d\mathscr{E} = \frac{1}{V}\int_{\mathscr{E}_c}^{\infty} g(\mathscr{E})\,\frac{1}{e^{\beta(\mathscr{E}-\mu)}+1}\, d\mathscr{E}, \qquad (5.66)$$

where $g(\mathscr{E})$ is the density of band states, of the type sketched in Fig. 5.64. The chemical potential lies somewhere in the gap between conduction and valence band

---

[22] Elements with different electron numbers, e.g. solid Na, Mg and Al, realize an energetically more stable configuration in more compact crystal structures (bcc, hexagonal, and fcc respectively), because their electrons are too few to fill completely the bonding-type band of a hypothetical diamond structure.

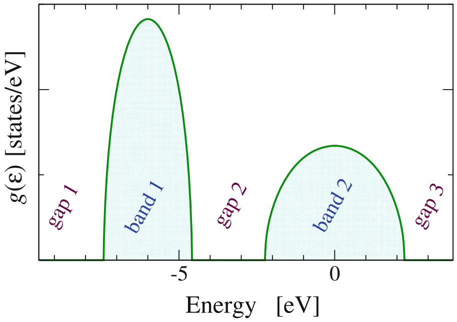


Fig. 5.64 A cartoon density of electronic states in a crystal. An actual solid usually displays narrow high-density bands at lower energy, and broader (often overlapping) low-density bands at higher energy. The band boundaries exhibit a characteristic $g(\mathscr{E}) \propto \left|\mathscr{E} - \mathscr{E}_{\text{boundary}}\right|^{1/2}$ energy dependence, a consequence of the quadratic $\mathbf{k}$ dependence of $\mathscr{E}_{\mathbf{k}\,j}$ near the band maximum or minimum [see Eq. (4.48)]. Gaps are characterized by $g(\mathscr{E}) = 0$

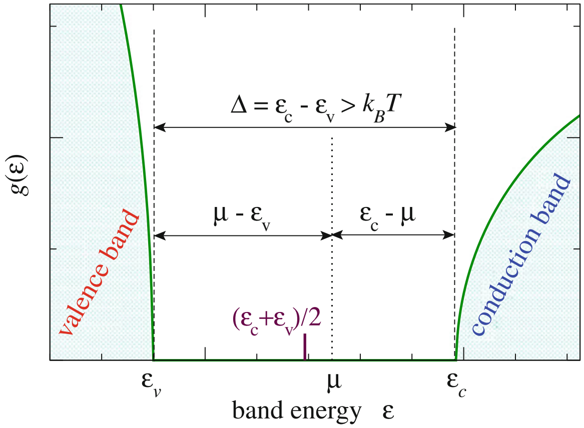


Fig. 5.65 At low temperature, in an intrinsic semiconductor characterized by an energy gap $\Delta = \mathscr{E}_c - \mathscr{E}_v$, the chemical potential μ lies near the mid-gap energy $(\mathscr{E}_c + \mathscr{E}_v)/2$. The energy distances of μ to the edges of both conduction and valence band are much larger than $k_{\mathrm{B}}T$. Accordingly, the electron occupation of individual conduction band states $[n_{\mathscr{E}_c}]_F$ is very small and that of valence states $[n_{\mathscr{E}_v}]_F$ is very close to unity

(at low temperature close to the mid-gap energy), thus several times $k_{\mathrm{B}}T$ below the conduction-band bottom $\mathscr{E}_c$ (see Fig. 5.65). It is therefore usually a good approximation to neglect the 1 at the denominator of $[n_{\mathscr{E}}]_F$, and take $(e^{\beta(\mathscr{E}-\mu)} + 1)^{-1} \simeq e^{-\beta(\mathscr{E}-\mu)} = e^{-\beta(\mathscr{E}_c-\mu)}\, e^{-\beta(\mathscr{E}-\mathscr{E}_c)}$. The first exponential is the same for all states in the band: it reflects the exponential suppression of the electron occupation of the conduction band due to its distance from μ. The second factor is the standard Boltzmann occupation of a high-temperature ideal gas, as in Eq. (4.10): fermion statistics has little or no effect in such an extremely rarefied electron gas. The energies $\mathscr{E}$ can be estimated by expanding the conduction band around its minimum[23]

$$\mathcal{E} = \mathcal{E}_{\mathbf{k}\,c} = \mathcal{E}_c + \frac{1}{2}\sum_{uw} \left.\frac{\partial^2 \mathcal{E}_{\mathbf{k}\,c}}{\partial k_u \partial k_w}\right|_{\mathbf{k}^{\min}} (k_u - k_u^{\min})(k_w - k_w^{\min}) + \cdots$$
$$\simeq \mathcal{E}_c + \frac{\hbar^2}{2m_c^*}|\mathbf{k} - \mathbf{k}^{\min}|^2 + \cdots$$

The approximate excitation energy $E = (\mathcal{E} - \mathcal{E}_c)$ in the second exponential takes the form of the kinetic energy of a free particle of mass m_c^*, once the wave numbers are referred to $\mathbf{k}^{\min}$. Accordingly, the density of conduction states (excluding spin) goes as $g_{\text{tr}}(E) = m_c^{*\,3/2} V/(\sqrt{2}\pi^2\hbar^3)\ E^{1/2}$ [see Eq. (4.48)]. The calculation of the $\mathcal{E}$ integration in Eq. (5.66) is then identical to the calculation of the classical partition function $Z_{1\,\text{tr}} = V/\Lambda_c^3$ of Eq. (4.42). In detail, we obtain

$$N_c = \frac{e^{-\beta(\mathcal{E}_c-\mu)}}{V}\int_{\mathcal{E}_c}^{\infty} g(\mathcal{E}) e^{-\beta(\mathcal{E}-\mathcal{E}_c)}\, d\mathcal{E} \simeq \frac{e^{-\beta(\mathcal{E}_c-\mu)}}{V}\int_0^{\infty} g_s\, g_{\text{tr}}(E)\, e^{-\beta E} dE$$
$$\simeq g_s \frac{e^{-\beta(\mathcal{E}_c-\mu)}}{V} Z_{1\,\text{tr}} = 2\frac{e^{-\beta(\mathcal{E}_c-\mu)}}{\Lambda_c^3} = 2\, e^{-\beta(\mathcal{E}_c-\mu)} \left(\frac{m_c^* k_{\text{B}} T}{2\pi\hbar^2}\right)^{3/2}, \qquad (5.67)$$

where the factor $g_s = 2$ reflects the spin degeneracy.[24]

The thermal length

$$\Lambda_c = \sqrt{\frac{2\pi\hbar^2}{m_c^* k_{\text{B}} T}} \simeq 6\,\text{nm},$$

at $T = 300\,\text{K}$ and taking $m_c^* \simeq 0.5\, m_e$. The exponential factor, for μ sitting at the middle of a 1.1 eV gap (i.e. $\mathcal{E}_c - \mu = 0.55\,\text{eV}$), is of the order $e^{\beta(\mu-\mathcal{E}_c)} \approx e^{-21} \simeq 6 \times 10^{-10}$. This yields about $N_c \approx 5 \times 10^{15}\,\text{m}^{-3}$, a modest charge-carrier density

---

[23] For energies high above the band minimum $\mathcal{E}_{\mathbf{k}\,c}$, the quadratic expansion is inaccurate, but the statistical occupation factor suppresses the contribution of the higher-energy states anyway.

[24] A very similar result is obtained for the number of *holes* in the valence band:

$$P_v = \frac{1}{V}\int_{-\infty}^{\mathcal{E}_v} g(\mathcal{E})(1 - [n_{\mathcal{E}}]_F)\, d\mathcal{E} \simeq e^{\beta(\mathcal{E}_v-\mu)} \frac{2}{\Lambda_v^3} = 2\, e^{\beta(\mathcal{E}_v-\mu)} \left(\frac{m_v^* k_{\text{B}} T}{2\pi\hbar^2}\right)^{3/2}.$$

The requirement of charge neutrality, $P_v = N_c$, fixes the position of the chemical potential: $e^{\beta(\mathcal{E}_v-\mu)}\, m_v^{*\,3/2} = e^{\beta(\mu-\mathcal{E}_c)}\, m_c^{*\,3/2}$, where we simplified common factors. By taking the logarithm of both sides we obtain

$$\mu = \frac{1}{2}(\mathcal{E}_v + \mathcal{E}_c) + \frac{3}{4} k_{\text{B}} T \ln \frac{m_v^*}{m_c^*},$$

which confirms that at $T = 0$ the chemical potential sits at the middle of the gap, and when T is raised it drifts slowly toward the band with the smaller effective mass, to compensate for the smaller density of states.

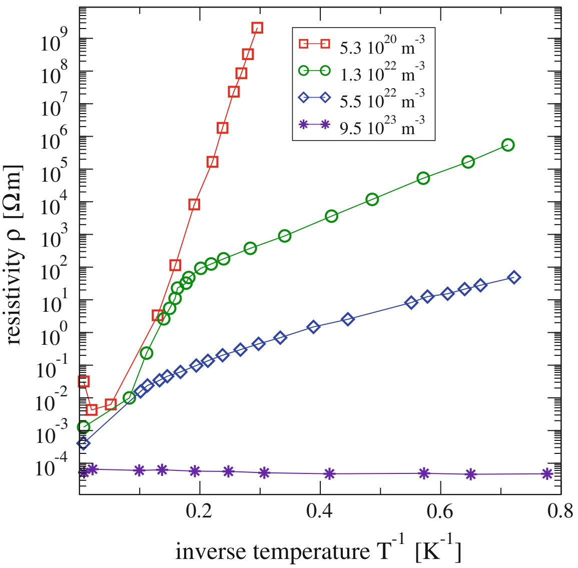


Fig. 5.66 The measured resistivity of antimony-doped germanium as a function of inverse temperature for increasing donor-impurity concentration, from $5.3 \times 10^{20}\,\mathrm{m}^{-3}$ (i.e. 0.012 part per million atoms) to $9.5 \times 10^{23}\,\mathrm{m}^{-3}$ (i.e. 21 part per million atoms) (Data from Ref. [40])

compared to that of regular metals ($\approx 10^{28}\,\mathrm{m}^{-3}$). Note however that this carrier density varies exponentially with T^{-1} (e.g. for the same conditions, $N_c \approx 6 \times 10^{20}\,\mathrm{m}^{-3}$ at $T = 600\,\mathrm{K}$). If this dependence is plugged into the expression (5.60) for conductivity in the presence of collisions, one expects a rather poor room-temperature conductivity, which is rapidly increasing with the exponential of T^{-1}, with slowly varying corrections due to (i) drifts of μ, (ii) the $\Lambda_c^{-3} \propto T^{3/2}$ term in Eq. (5.67), and (iii) a decreasing τ due to the increasing collision rate with phonons. Indeed, intrinsic semiconductors (square points in Fig. 5.66) exhibit a room-temperature resistivity many orders of magnitude larger than in metals (Table 5.2). The drastic temperature dependence (approximately exponential in T^{-1}) of resistivity makes pure semiconductors quite sensitive temperature sensors, especially at in the few-K range where the resistivity of metals is almost temperature-independent—see Fig. 5.59.

In practice, semiconductors find important applications mainly as "doped" crystals, called *extrinsic semiconductors*. Doping is realized typically by substitutional impurities replacing a few of the perfect-crystal atoms, as sketched in Fig. 5.67. Pentavalent impurities, such as P or As replace Si/Ge atoms at the regular lattice sites, thus formally establishing four chemical bonds. In other terms, the crystal bands are not much deformed by the impurities. However a pentavalent atom carries one extra positive nuclear charge, thus forming a potential well which tends to attract one extra electron. This well produces a characteristic localized "impurity" state, associated to an energy level inside the band gap. At zero temperature, the extra electron carried by the pentavalent atom occupies this impurity state.

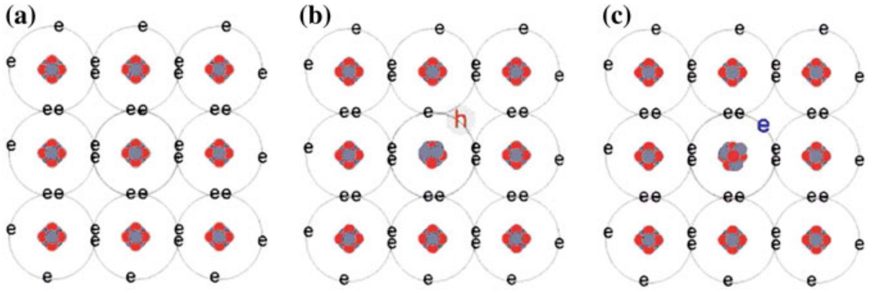


Fig. 5.67 Substitutional atoms of group III or group V replace a few Si or Ge atoms of the pure semiconductor (**a**), producing extrinsic semiconductors of p type (**b**) or n type (**c**) respectively. The square lattice is just a convenient pictorial for the actual diamond structure

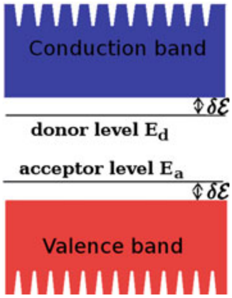


Fig. 5.68 Localized donor and acceptor levels are very shallow. They are usually located within few tens meV of the conduction-band minimum and valence-band maximum respectively. Room temperature is hot enough to excite carriers from these localized states to the delocalized band states

The main feature of the impurity states of pentavalent dopants (*donors*) is their close vicinity to the conduction band (Fig. 5.68). Due to screening, the extra electron is bound to the impurity ion very weakly. The electron binding to the impurity can be roughly described as a particle of charge $-q_e$ and mass m_c^* attracted to the impurity nucleus by the screened Coulomb potential of the impurity ion $\simeq q_e/(4\pi\varepsilon_0\varepsilon\, r)$, where ε is the static dielectric constant of the pure semiconductor ($\varepsilon \simeq 12$ for Si, $\varepsilon \simeq 16$ for Ge). The bound-state energy levels of this 1-electron-atom-like model are given by Eq. (2.10) with effective nuclear charge $Z = \varepsilon^{-1}$ and effective mass $\mu = m_c^*$. In particular, the ground-state energy equals a suitably-rescaled Rydberg energy:

$$\delta\mathscr{E} \simeq \frac{m_c^*}{m_e}\,\frac{1}{\varepsilon^2}\,\frac{E_{\mathrm{Ha}}}{2}. \tag{5.68}$$

Typical values of m_c^* and ε lead to binding energies of the order $10^{-3}E_{\mathrm{Ha}} \simeq 30\,\mathrm{meV}$. Indeed, the separations $\delta\mathscr{E} = \mathscr{E}_c - E_d$ of the P and As impurity levels from the bottom of the conduction band of Si are observed 44 and 49 meV respectively (in Ge, 12 and 13 meV respectively).

A trivalent impurity (*acceptor*) carries a localized excess negative charge, as long as the conduction band is filled. The missing electron can be represented as a bound *hole*, weakly attracted to the excess negative charge representing the impurity. In

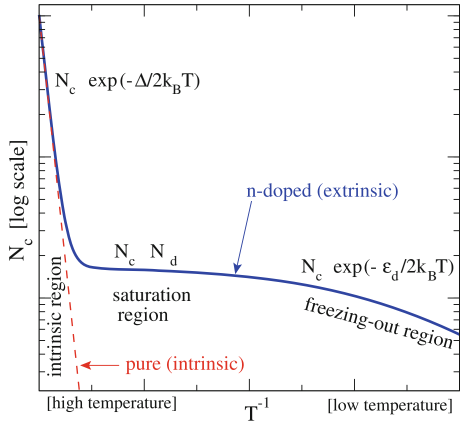


Fig. 5.69 A cartoon temperature dependence of the majority carrier density for n doping. The high-temperature regime is dominated by intrinsic carriers. The intermediate "saturation" regime of almost constant $N_c \simeq N_d$ is the temperature region where extrinsic carriers prevail and are dissociated from their impurities. The low-temperature decrease of N_c is due to "freezing" of the extrinsic carriers, which at absolute zero would all end up "captured" by the localized impurity states

the electron picture this bound hole will manifested itself as an additional electronic level E_a slightly above the top of the valence band. The hole is bound when this level is empty, as at $T = 0$. When an electron is promoted from the valence band into this localized level, paying a small energy $E_a - E_v$, the excess charge of the impurity is removed, and an unbound hole is left in the valence band.

Impurity states are localized and do not contribute to transport. At $T = 0$, a homogeneous doped semiconductor is an insulator, with the Fermi level sitting near either E_a or E_d, according to whether the density N_a of acceptors or that N_d of donors is larger (*p/n doping* respectively). As temperature is raised from 0, the bound charges get rapidly unbound into the band levels, mostly in the valence band, if $N_a > N_d$ (p doping), or in the conduction band, if $N_d > N_a$ (n doping). Due to the small binding energy of the impurity levels compared to the band gap, it is far easier to thermally excite an electron into the conduction band from a donor level, or a hole into the valence band from an acceptor level, than it is to excite an electron across the entire gap Δ from valence to conduction band. At room temperature, the probability that an electron unbinds from the impurity levels is high. The chemical potential moves toward the middle of the gap, but (for temperature not too high) it remains closer to the valence (p doping) or conduction (n doping) band. This leads to a significant concentration of carriers, which dominates over the intrinsic carriers over a broad temperature range. In this "saturation" regime the density of majority carriers (holes for p doping, electrons for n doping) changes slowly and approximates the net concentration of impurities,[25] as in Fig. 5.69. This carrier concentration, plugged into

---

[25] For n doping an electron density $N_c \simeq N_d - N_a$. For p doping a hole density $P_v \simeq N_a - N_d$.

Eq. (5.60), is compatible with the doping and temperature dependence of resistivity shown in Fig. 5.66. In the saturation regime, doped semiconductors conduct like poor metals. In particular, p-doped semiconductors exhibit reversed Hall field, indicating hole conduction. The carrier population in the minority band (conduction for p doping and valence for n doping) is extremely small (but increases rapidly with temperature).

Doped semiconductors deliver an immense range of applications mostly as inhomogeneous systems, i.e. crystals where the impurity concentrations vary in space. Advanced techniques allow industries and labs to tailor the doping level over a sub-μm scale, across crystals of lateral size exceeding several centimeters. This technology is at the basis of the modern electronics industry, where semiconductor devices act as “active” components which permit the manipulation (e.g. amplification or shaping) of electric signals. Countless other applications of inhomogeneous semiconductor include sensors, light production, electronic data manipulation… Since the 1950s, semiconductor devices have been a leading area of research and development, a common playground of fundamental QM, solid-state physics, materials science, industrial engineering, and electronics. Specific courses delve in this vast field. Here we only sketch the principle of functioning of the simplest inhomogeneous extrinsic semiconductor: the *p-n junction*.

Consider a piece of semiconductor with ideal step-like p-n dopant densities

$$N_a(x) = \begin{cases} N_a, & x < 0 \\ 0, & x > 0 \end{cases}$$
$$N_d(x) = \begin{cases} 0, & x < 0 \\ N_d, & x > 0 \end{cases}, \tag{5.69}$$

as a function of some displacement x across the sample. This is what one could conceptually (not practically) construct by assembling a p-doped and a n-doped piece of semiconductor. As required by thermodynamic equilibrium, the chemical potential in the two separate sections of the semiconductor, initially significantly different, must become the same after this contact is realized. Starting with each portion in a homogeneous neutral situation, the equilibration of the chemical potential is realized by the diffusion of electrons from the n side (where their concentration is higher) into the p side (where their concentration is smaller), and of holes in the opposite direction, like when a wall separating oxygen and nitrogen is removed from the middle of a vessel and the two gases mix. As diffusion continues, the resulting charge transfer builds up an electric field opposing further diffusive currents until a steady configuration is reached, when the electric force on the charge carriers balances the effect of diffusion.

It is possible to describe the electric field $E_x = -d\phi/dx$ and the local densities of electrons and holes *via* coupled Maxwell equations and Eqs. (5.54) and (5.55) for the carriers. However, here we adopt a simpler qualitative approach. Because the carriers are highly mobile, in the equilibrium configuration the carrier densities are very low wherever E_x has an appreciable value, see Fig. 5.70b. The resulting *depletion layer* is typically of 10–10^3 nm thick, depending mainly on the dopant

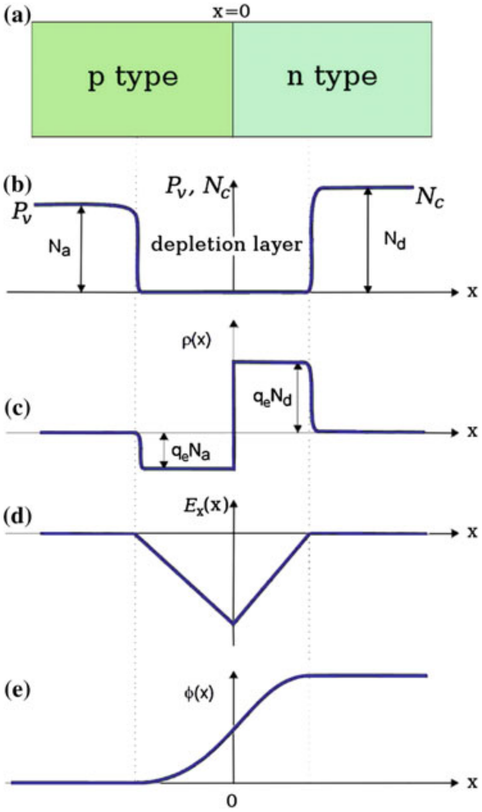


Fig. 5.70 **a** A sharp p-n junction at thermodynamic equilibrium (no voltage bias). **b** The qualitative profile of the hole density $P_v(x)$ (*left*) and electron density $N_c(x)$ (*right*). These quantities approach rapidly the bulk values $P_v(x \ll 0) \simeq N_a$, $N_c(x \gg 0) \simeq N_d$ that guarantee charge neutrality far away from the junction. In the region near the junction, both carrier concentrations are strongly suppressed: hence the name *depletion region*. **c** The total electric charge density $\rho(x)$: the uncompensated densities N_a and N_d of acceptors and donors dominate a double layer, leaving a net negative charge at the p side of the junction and a net positive charge at its n side. **d** This double layer of charge produces a leftward electric field $E_x(x)$, which sweeps the carriers away, maintaining the depletion. **e** The electric potential $\phi(x)$ consistent with $E_x(x)$

concentrations and the dielectric constant of the semiconductor. The impurity-ion charges remain uncompensated in the depletion layer, thus producing the electric charge-density profile illustrated in Fig. 5.70c. The double layer of charge generates the electric field sketched in Fig. 5.70d. This field corresponds to a finite potential drop (Fig. 5.70e), which compensates the different chemical potential in the bulk, far away from the junction:

$$q_e \Delta\phi_0 = q_e[\phi(+\infty) - \phi(-\infty)] = \mu_{\rm n} - \mu_{\rm p}, \tag{5.70}$$

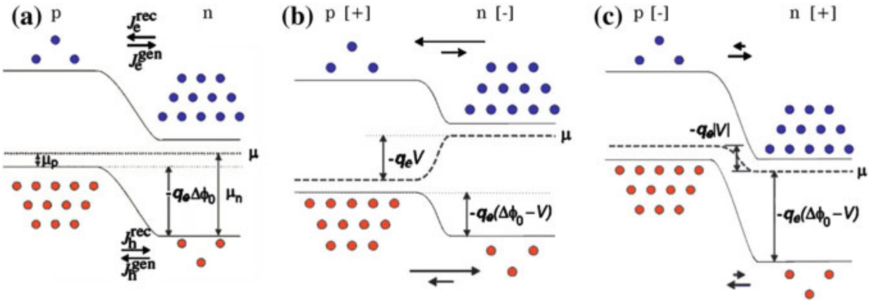


Fig. 5.71 The energy barrier $-q_e\phi(x)$ encountered by negatively-charged electrons, as a function of the displacement x across the p-n junction. **a** At equilibrium, in the absence of external bias. **b** In the presence of forward bias ($V > 0$). **c** In the presence of reverse bias ($V < 0$). Carrier currents are indicated: note that generation currents are largely independent of bias

where $\mu_{\mathrm{p/n}}$ indicate the chemical potential in the isolated bulk p or n semiconductors, prior to the construction of the junction. This potential difference shifts rigidly the bands (and impurity levels) away from the junction, as illustrated in Fig. 5.71a. When metal electrodes are deposited on both sides of the junction and joined by a wire, at thermodynamic equilibrium the chemical potential aligns everywhere inside the circuit to a common value, and no net current flows in the circuit, as the individual potential drops at the interfaces cancel algebraically.

At the p-n junction the net current vanishes due to the cancellation of four contributions:

- At the p side of the depletion layer a few "minority" electrons appear by thermal excitation out of the valence band: these are immediately "swept" by the strong electric field, and expelled to the n side (Fig. 5.71a). This electron *generation current* J_e^{gen} depends exponentially on temperature, but only weakly on the potential drop and the size of the depletion region.
- Few (majority) electrons at the n side of the junction acquire enough thermal energy to overcome the potential barrier to the p side. Once in the p region, these electrons turn into minority carriers and are likely to recombine with (majority) holes in the valence band: the corresponding current is therefore called *recombination current* J_e^{rec}.
- Similarly, a hole generation current J_h^{gen} sweeps thermal holes from the n to the p side, and a hole recombination current J_h^{rec} accounts for thermally excited holes diffusing into the n side.

At equilibrium these currents cancel in pairs ($J_e^{\mathrm{gen}} = J_e^{\mathrm{rec}}$, and $J_h^{\mathrm{gen}} = J_h^{\mathrm{rec}}$) and no net current flows through the junction.[26]

The external shorting wire can be cut and a voltage generator inserted, to alter the electric potential drop across the semiconductor by a tunable potential V, with the

---

[26] The four quantities $J_{e/h}^{\mathrm{rec/gen}}$ are defined as the (positive) norm of the corresponding number current-density vectors, and are measured in units of $\mathrm{s}^{-1}\,\mathrm{m}^{-2}$.

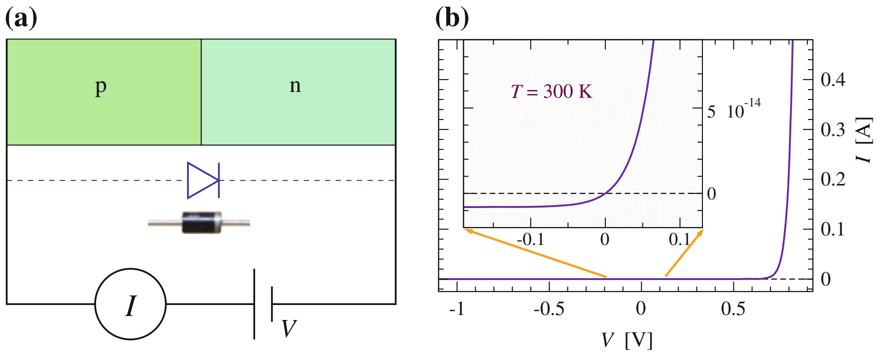


Fig. 5.72 **a** Circuit scheme for measuring the $I - V$ characteristics of a p-n junction. The sign convention has forward bias $V > 0$ increasing the electric potential at the p side and producing a fast-rising current $I > 0$; $V < 0$ (reverse bias) sustains a tiny essentially V-independent reverse current $I < 0$. The diode electronic symbol plus a packaged silicon diode are also depicted aside the p-n junction. **b** The $I - V$ characteristic according to Eq. (5.73)

sign convention of Fig. 5.72. The bulk p and n regions are fair (low-resistance) conductors, thanks to their large carrier density (compared to the depletion layer). We can therefore assume that practically all the potential drop induced by the external applied field is realized over the depletion layer, as illustrated in Fig. 5.71. When $V \neq 0$ the condition of thermodynamic equilibrium is violated and a net current density j_x (thus a current I) is established through the junction. To understand semi-quantitatively the $I - V$ characteristic of the p-n junction, observe that the recombination currents depend very strongly on the electric potential drop through the depletion layer: J_e^{rec} is proportional to the number of carriers acquiring sufficient energy to overcome the potential barrier. This is now modified by V: $J_e^{\text{rec}} \propto \exp[-q_e(\Delta\phi_0 - V)/(k_\text{B}T)]$. The equilibrium requirement at $V = 0$ ($J_e^{\text{rec}} = J_e^{\text{gen}}$) and the fact that the generation current is almost independent of V fixes the proportionality constant:

$$J_e^{\text{rec}} = J_e^{\text{gen}}\, e^{\frac{q_e V}{k_\text{B}T}}. \tag{5.71}$$

The total current density carried by electrons is therefore

$$J_e = J_e^{\text{rec}} - J_e^{\text{gen}} = J_e^{\text{gen}} \left(e^{\frac{q_e V}{k_\text{B}T}} - 1 \right). \tag{5.72}$$

A similar analysis, with a similar result, can be carried out for holes. Compared to electrons, the hole currents J_h^{gen} and J_h^{rec} move in the opposite direction, but carry positive charge, thus their contribution adds up to that of electrons, to give a total electric current density

$$j_x = q_e \left(J_e^{\text{gen}} + J_h^{\text{gen}} \right) \left(e^{\frac{q_e V}{k_\text{B}T}} - 1 \right). \tag{5.73}$$

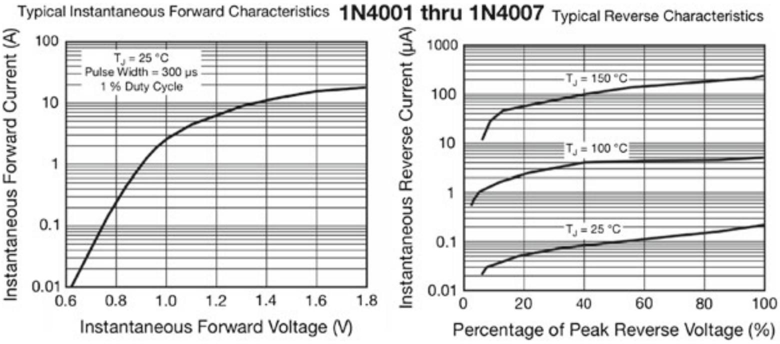


Fig. 5.73 Typical $I - V$ characteristics of commercial silicon diodes: Vishay 1N4001–1N4007. Observe deviations from the ideal behavior of Eq. (5.73): (i) The exponential current increase as a function of forward bias V ends at about $V = 0.8$ V; above this bias the current is limited by the series resistance of the homogeneous regions of the doped semiconductor. (ii) The reverse current is significantly larger than predicted by Eq. (5.73) and keeps changing slowly with V instead of remaining constant for $-V > 4k_{\mathrm{B}}T/q_e \simeq 0.1$ V. (iii) A reverse breakdown knee is observed as the reverse voltage exceeds the rated peak reverse voltage (50–1000 V for this family of diodes). (From the web site http://www.diotec-usa.com/; copyright 2010 by Diotec)

This $I - V$ characteristic, illustrated in Fig. 5.72, is strongly asymmetric and non-linear: basically the p-n junction operates as a rectifier, allowing electric current to circulate in one direction only, similarly to the old vacuum diode based on thermionic emission. For this reason, a two-terminal semiconductor device consisting of a single p-n junction is named a *diode*. $I - V$ curves very much like Eq. (5.73) are indeed observed experimentally (Fig. 5.73). Deviations include a reverse breakdown regime at large negative voltage, and resistive current limitation at large positive voltage.

Carefully tailored commercial p-n junctions are built for all sorts of applications. They include: the light emitting diode (LED, where forward electrical bias causes the two species of charge carrier—holes and electrons—to be “injected” into the depletion region, where their spontaneous recombination generates visible or IR light depending on the semiconductor gap), the laser diode (like a LED plus an optical cavity surrounding the p-n junction to enhance stimulated emission), the photo-voltaic cell (the reverse of a LED: light induces extra electron-hole pairs in the depletion layer, and these carriers drift following the local field and reduce the junction electric potential drop from its equilibrium value $\Delta\phi_0$, thus inducing a current in an external circuit, and converting radiation into electric power).

Beside 2-terminal devices, semiconductor single crystals are doped to form several p-n junctions. The simplest 3-terminal device is the bipolar *transistor*, a two-junctions sandwich, where the voltage (and a small current) across one junction controls the (potentially large) current through the second junction, thus providing amplification. Amplification is also realized by the constructively different field-emission transistor (FET). Diodes, transistors, and other semiconductor electronic components provide the ability to manipulate electronic signals. These components (often thousands or even millions of them) are routinely integrated into large semiconductor (usually silicon) single crystals of $\sim$cm lateral size, to provide the complex functionality at the core of modern microelectronics [41].

5.2.3 *Spectra of Electrons in Solids*

The experimental investigation of the electronic states in solids relies on several techniques. Here we discuss briefly three basic ones: photoemission, X-ray absorption/emission, and optical absorption/reflectivity.

Figure 5.74 illustrates the conceptual setup of a *photoemission* experiment. The kinetic-energy $E_{\rm kin}$ distribution $I(E_{\rm kin})$ of the photoemitted electrons measures the density of *occupied* states of the material, with a rigid shift by the photon energy $\hbar\omega$. The observed continuous spectrum probes the filled bands of the material, essentially up to the chemical potential. Angle-integrated photoemission gives direct information on the density of occupied band states $g(\mathscr{E})[n_{\mathscr{E}}]_F$, and on forbidden energy gaps, as in Fig. 5.75. Moreover, in metals, photoemission in the region $\mathscr{E} \simeq \mu$ probes the Fermi edge, i.e. the step-like drop of $[n_{\mathscr{E}}]_F$ close to μ. Angle-resolved photoemission provides further information about the $\mathbf{k}$-dependence of the band energy $\mathscr{E}_{\mathbf{k}\,j}$. A possible source of systematic error in photoemission is the charge buildup in insu-

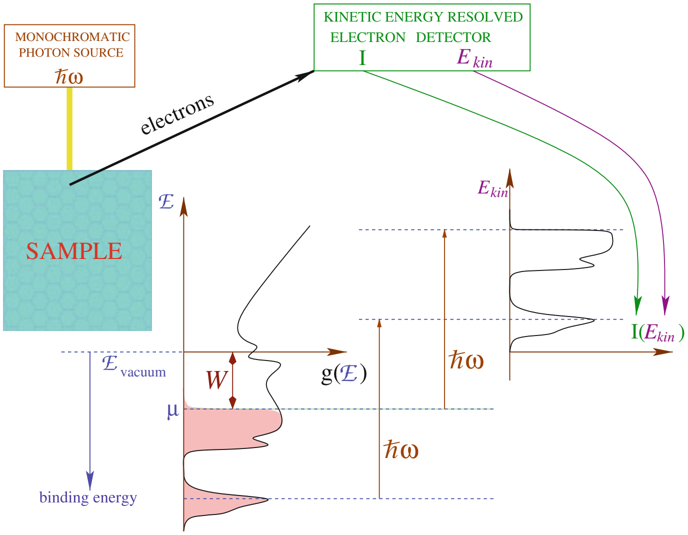


Fig. 5.74 The scheme and interpretation of a photoemission experiment. An electron initially in a band state of the solid absorbs a photon and acquires its energy $\hbar\omega$. If the energy of the resulting final state is sufficient, the electron can move out of the solid into the surrounding vacuum. The detector measures the kinetic energy $E_{\rm kin} = \mathscr{E} + \hbar\omega$ distribution $I(E_{\rm kin})$ of the photoemitted electrons thus mapping the density of *occupied* states of the sample, up to the Fermi energy

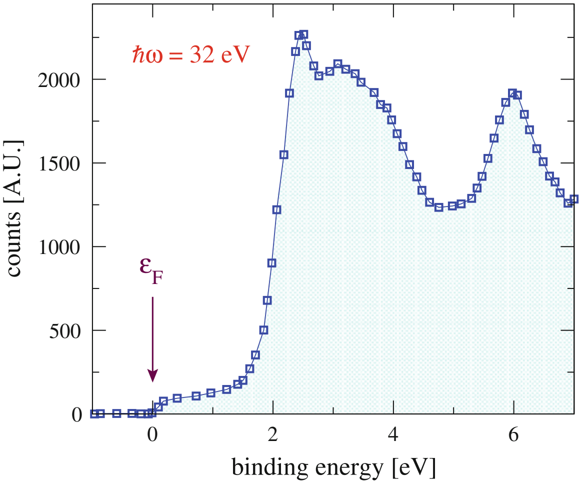


Fig. 5.75 The observed photoemission spectrum of a gold film. "Binding energy" should indicate $\mathscr{E} = (\hbar\omega - E_{\text{kin}})$, as in the left scale of Fig. 4.13. However, as is common practice, here the scale zero is shifted to the Fermi level, rather than at the vacuum level. The weak intensity between 0 and 1.5 eV represents the broad metallic (mainly 6s) band; the intense structures in the 2–7 eV range are associated to the narrow 5d bands

lating samples caused by electron removal and producing a macroscopic electric field outside the sample. Further factors limiting the quality of the spectra are: secondary collisions of the emitted electron while still inside the solid, and the short penetration of visible/UV light in metals. Both these effects make photoemission more sensitive to surface layers than to the solid bulk.

These difficulties are circumvented by X-ray spectroscopies. Resonant *X-ray absorption*, in particular, excites core electrons into the empty bands states, thus probing their density of states $g(\mathscr{E})(1 - [n_{\mathscr{E}}]_F)$, see Figs. 5.76 and 5.77a.[27] *X-ray emission*, instead, probes the density of filled band states $g(\mathscr{E})[n_{\mathscr{E}}]_F$, whence elec-

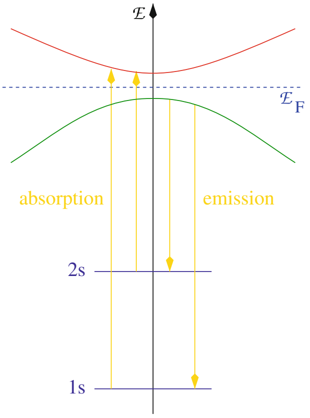


Fig. 5.76 An elementary interpretation of X-ray absorption and emission spectra

---

[27] The interpretation of these spectra is complicated by the presence of the core hole, which acts as a localized extra charge and distorts the bands as if an impurity was located at that site.

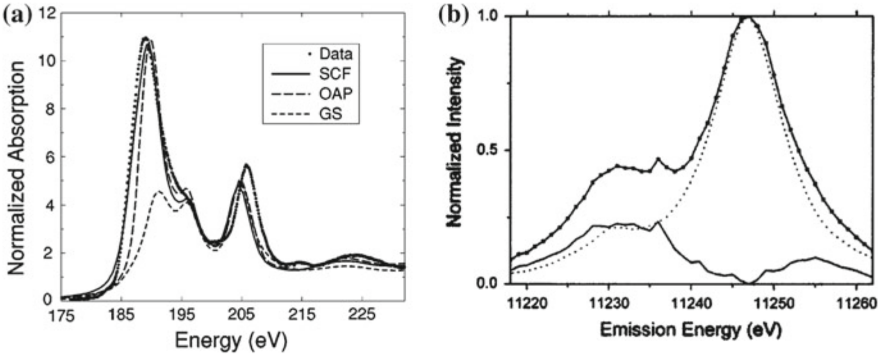


Fig. 5.77 **a** *Points* the X-ray absorption spectrum of cubic boron nitride (BN) at the B 1s edge, probing the empty band levels of this material. *Continuous curves* model calculations which interpret the experimental data. Reprinted figure with permission from A.L. Ankudinov, B. Ravel, J.J. Rehr, and S.D. Conradson, Phys. Rev. B **58**, 7565 (1998). Copyright (1998) by the American Physical Society. **b** The 4d→2$p_{3/2}$ and 3p→2s emission spectra of Pt excited at respectively 11,610 eV (*dotted*) and 14,010 eV (*connected points*), probing the density of occupied states [42]. Reprinted figure with permission from F.M.F. de Groot, M.H. Krisch, and J. Vogel, Phys. Rev. B **66**, 195112 (2002). Copyright (2002) by the American Physical Society

trons can decay to fill a core hole previously generated (Fig. 5.77b). The main disadvantage of core spectroscopies is the severe limitation in energy resolution imposed by the broad short-lived core-hole states (Sects. 1.2 and 2.2.7).

Optical excitations are usually probed by means of optical *reflectivity*, rather than absorption, since solids are often too opaque to transmit significant visible or UV radiation across any practical sample thickness, see Fig. 5.78. Both reflectivity and absorption are related to the material's complex dielectric function $\varepsilon(\omega)$, that accounts for both the change of speed of radiation and the energy transferred from the radiation field to the material. At low frequency, the reflectivity of metals approaches 100 % due to intra-band electron excitations, see Fig. 5.78.[28] At larger energy, of the order of few to several eV, inter-band transitions dominate. Different colors of different metals (Cu, Ag, Au, Fe) indicate different spectral reflectivities, which can be understood in terms of their different band structures.

Inter-band transitions conserve energy and momentum (thus wave number):

---

[28] In the limit of extremely small photon energy (long wavelength), the DC regime is approached, where a good metal "short circuits" the electric-field component of radiation, thus efficiently suppressing it inside the material, and reflecting it outside. Note that reflectivity is defined for an infinitely thick metal layer. Note also that the absorption coefficient α decreases slowly as ω decreases down into the radio-wave region. As a consequence of these two observation, in practice radio-waves are not 100 % reflected, but partly penetrate through a sufficiently thin metal layer, of thickness not exceeding a few times α^{-1}. This is the reason why one can operate a cell phone e.g. inside a metallic elevator cab.

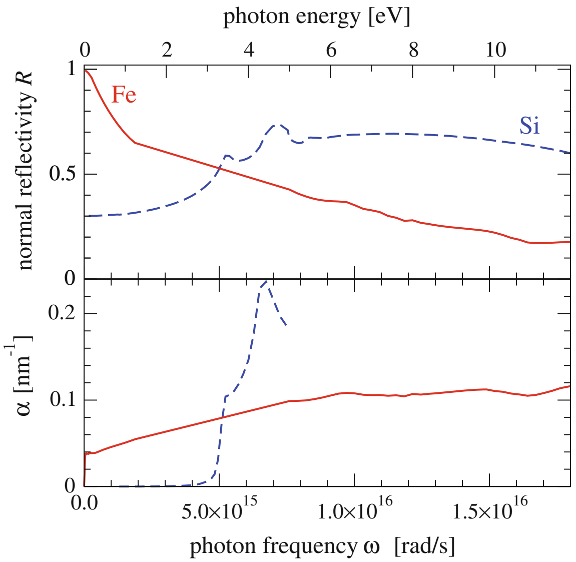


Fig. 5.78 The reflectivity R at normal incidence and the absorption coefficient α of a metal (iron) and a semiconductor (silicon) at room temperature, for electromagnetic radiation around the optical region. In the static ($\omega \to 0$) limit, the reflectivity R of metals approaches unity, as characteristic of intra-band transitions. In the same limit, the reflectivity of semiconductors and insulator approaches a constant value connected to their static dielectric permittivity by $R = (\varepsilon^{1/2}-1)^2/(\varepsilon^{1/2}+1)^2 \simeq 30\,\%$ for Si, which has $\varepsilon \simeq 12$. The wiggles in the visible-UV region originate from inter-band transitions. In the far-UV and X-ray region both R and α decay smoothly toward 0, except for specific structures at the core-shell energies, see Fig. 2.22 (Fe data from Ref. [43], Si data from Ref. [44])

$$\hbar\omega = \mathscr{E}_{\mathbf{k}'\,j'} - \mathscr{E}_{\mathbf{k}\,j}, \quad \mathbf{q} = \mathbf{k}' - \mathbf{k}, \tag{5.74}$$

where $\hbar\omega$ is the photon energy, $\mathbf{q}$ is the photon wave number and, as usual in a crystal, the initial and final $\mathbf{k}$ and $\mathbf{k}'$ are defined up to a reciprocal lattice vector $\mathbf{G}$, and can therefore be reported within the first BZ. Importantly, at optical wavelengths $\lambda \sim 500\,\text{nm}$, the photon wave number $|\mathbf{q}| \approx 10^7\,\text{m}^{-1}$, while typical wave numbers of electrons in a solid are of the order of the size of the first BZ, i.e. of the order of the reciprocal lattice parameter, in the 10^9–$10^{10}\,\text{m}^{-1}$ region. Therefore in the conservation (5.74) the photon wave number $\mathbf{q}$ can be neglected, implying $\mathbf{k}' \simeq \mathbf{k}$. In words, optical transitions are "vertical" in $\mathbf{k}$ space, as sketched in Fig. 5.79.[29]

Unlike metals, pure semiconductors and insulators have no intra-band excitations, and are therefore basically transparent at low-frequency, with very weak absorption up to photon energies at least as large as the gap Δ. In direct-gap semiconductors, the

---

[29] In insulators and intrinsic semiconductors, the conduction electrons and valence holes induced by photon absorption can carry current as if they were generated by high temperature or by doping: this phenomenon, named *photoconductivity*, is exploited in light sensors.

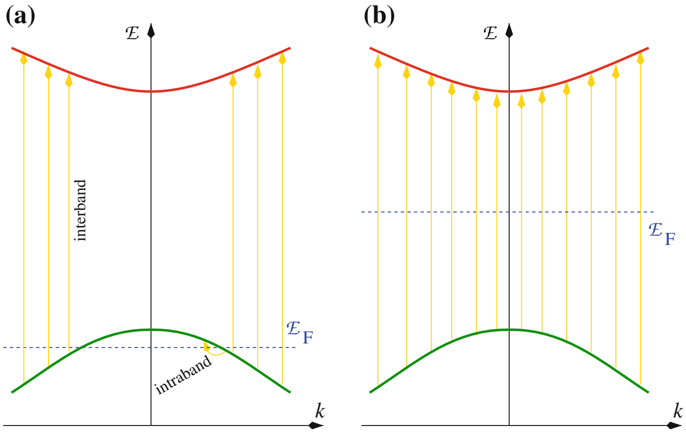


Fig. 5.79 A sketch of the excitations probed in optical absorption/reflectivity experiments of crystalline solids. **a** A metal; **b** a direct-gap insulator. Transitions are vertical due to wave-vector conservation

absorption threshold measures the gap amplitude directly. However, when the gap is indirect (see Fig. 5.63), the onset of vertical transitions occurs at an energy larger than the actual gap. This difference is quite large for silicon, which has $\Delta \simeq 1.17\,\text{eV}$ but little absorption before $\sim 3\,\text{eV}$, see Fig. 5.78. With the assistance of a phonon (of much smaller energy than a typical electronic energy) providing the missing wave vector, the smallest-energy nonvertical transition across the gap can then be observed in an "indirect" way.[30]

This simplified picture relating optical properties to the electronic band structure relies on the assumption that photoexcited electrons and holes behave as independent (quasi-)particles propagating in the conduction and valence bands, respectively. While this approximation may work in semiconductors, in wide-gap insulators screening by the other electrons is small and the attraction between the electron and the hole cannot be neglected. This electron-hole interaction can modify strongly the optical absorption near and below the gap [12, 45], often creating bound electron-hole pairs ("excitons"). This characteristic effect is viewed e.g. in the UV spectrum of LiF, Fig. 5.80, which exhibits a prominent excitonic peak well below the gap.

Impurities in otherwise perfect insulators provide localized states inside the gap: transitions involving these impurity states often appear with strong intensity. For example, perfect industrial grade Al_2O_3 (corundum) is transparent to visible light ($\Delta \approx 5\,\text{eV}$). However Al_2O_3 scatters characteristic red or blue light upon Cr or Ti doping: in these doped forms Al_2O_3 is commonly called "ruby" or "sapphire" respectively in the gemstone trade, see Fig. 5.10c, d.

---

[30] The need to involve a phonon makes these transitions far less likely: for this reason, semiconductor applications involving light detection/production/control (optoelectronics) usually prefer direct-gap semiconductors.

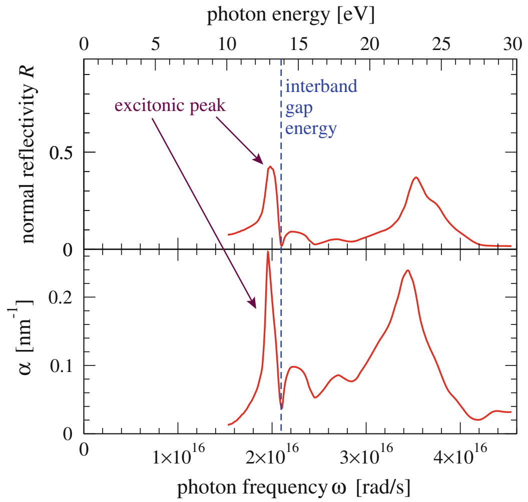


Fig. 5.80 Same as Fig. 5.78, but for the wide-bandgap insulator LiF. The intense peak below the bandgap width (*dashed line*) is attributed to an excitonic state, i.e. an electron-hole bound state (Data from Ref. [46])

5.3 The Vibrations of Crystals

In Sect. 5.2 the atomic nuclei were assumed to sit quietly at their more or less ideal crystalline equilibrium positions. This static-ion model is fine for describing many properties of solids. However, several other observed properties including elasticity, heat capacity, and heat transport depend on the vibrational dynamics and thermodynamics of ions. The reader will note a similarity with the effects of the internal motions of molecular gases—Sects. 3.3 and 4.3.1.2.

5.3.1 The Normal Modes of Vibration

Let us go back to the Eq. (3.9) for the motion of the nuclei. We assume that the adiabatic potential $V_{\mathrm{ad}}(R)$ of Eq. (3.10) has a single well-defined minimum.[31] Like we did for molecules in Eq. (3.20), we expand $V_{\mathrm{ad}}(R)$ around its minimum R_{M}, recalling that the first-order term proportional to gradient of $V_{\mathrm{ad}}(R)$ vanishes at R_{M} because the minimum is a stationary point:

---

[31] For many materials, a unique absolute minimum R_{M} of $V_{\mathrm{ad}}(R)$ represents the perfect crystal. However, the actual configuration of a solid often deviates substantially, as it remains trapped in a defective configuration, represented by one of many deep local minima occurring in the configurations space somewhere around R_{M}. The thermal and quantum tunneling rate through the barrier leading from the defective configuration to the perfect crystal is often negligible, and the defected configuration is metastable. The vibrations around a moderately defective local minimum might not differ too much from those around the crystalline absolute minimum.

$$V_{\text{ad}}(R) = V_{\text{ad}}(R_{\text{M}}) + \frac{1}{2}\sum_{\alpha\beta} \left.\frac{\partial^2 V_{\text{ad}}}{\partial R_\alpha \partial R_\beta}\right|_{R_{\text{M}}} u_\alpha u_\beta + \mathcal{O}(u^3), \qquad (5.75)$$

where the vector u of components $u_\alpha = R_\alpha - R_{\text{M}\,\alpha}$ collects the atomic displacements away from the minimum

$$\begin{pmatrix} u_{1x} \\ u_{1y} \\ u_{1z} \\ u_{2x} \\ u_{2y} \\ u_{2z} \\ u_{3x} \\ \vdots \\ u_{N_n x} \\ u_{N_n y} \\ u_{N_n z} \end{pmatrix} = \begin{pmatrix} (R)_{1x} \\ (R)_{1y} \\ (R)_{1z} \\ (R)_{2x} \\ (R)_{2y} \\ (R)_{2z} \\ (R)_{3x} \\ \vdots \\ (R)_{N_n x} \\ (R)_{N_n y} \\ (R)_{N_n z} \end{pmatrix} - \begin{pmatrix} (R_{\text{M}})_{1x} \\ (R_{\text{M}})_{1y} \\ (R_{\text{M}})_{1z} \\ (R_{\text{M}})_{2x} \\ (R_{\text{M}})_{2y} \\ (R_{\text{M}})_{2z} \\ (R_{\text{M}})_{3x} \\ \vdots \\ (R_{\text{M}})_{N_n x} \\ (R_{\text{M}})_{N_n y} \\ (R_{\text{M}})_{N_n z} \end{pmatrix}. \qquad (5.76)$$

The atomic nuclei are labeled $j = 1, 2, \ldots, N_n$, with component $\xi = x, y, z$; α stands collectively for j, ξ, and thus spans $3N_n$ values. Call Φ the $3N_n \times 3N_n$ (Hessian) matrix of second derivatives of $V_{\text{ad}}(R)$: the element $\Phi_{\alpha\alpha'} = \Phi_{j\xi\; j'\xi'}$ in the second-order term of the expansion (5.75) amounts to minus the ξ-change in the force acting on atom j due to a unit displacement of atom j' in direction ξ'. By neglecting all terms of third and higher order in u in Eq. (5.75) we obtain the *harmonic approximation*. This approximation is meaningful to describe small oscillations around equilibrium, thus it is usually appropriate for solids at low temperature, but it is bound to fail e.g. when the solid is highly deformed, or near melting.

The Schrödinger equation (3.9) could be solved exactly for a quadratic multidimensional potential. Unfortunately, this fully quantum approach involves some mathematical intricacies. It is preferable to first solve the *classical* Newton-Hamilton equations for the motion of the ions, and introduce QM as a later step. In vector form, the $3N_n$ classical equations of motion for the u dynamical variables are:

$$\begin{aligned} M_n \frac{d^2 u}{dt^2} &= -\nabla_u V_{\text{ad}}(R_{\text{M}} + u) \\ &\simeq -\nabla_u \left[V_{\text{ad}}(R_{\text{M}}) + \frac{1}{2}\sum_{\alpha'\beta'} \Phi_{\alpha'\beta'} u_{\alpha'} u_{\beta'} \right] = -\frac{1}{2}\nabla_u \sum_{\alpha'\beta'} \Phi_{\alpha'\beta'} u_{\alpha'} u_{\beta'}, \end{aligned} \qquad (5.77)$$

(assume for simplicity all atoms of the same mass M_n). Equation (5.77) can be written for a generic α-component of u, by noting that u_α appears twice in the double sum, as either $u_{\alpha'}$ or $u_{\beta'}$, thus leading to a single sum and canceling the factor 1/2:

$$M_n \frac{d^2 u_\alpha}{dt^2} = -\sum_\beta \Phi_{\alpha\beta}\, u_\beta. \tag{5.78}$$

We have obtained a set of $3N_n$ equations for $3N_n$ coupled harmonic oscillators. Couplings are introduced by the off-diagonal matrix elements of Φ, i.e. by the forces that the displacement of an atom produces on other atoms.

A *normal mode* is a periodic motion of all atoms oscillating at a single frequency. To investigate normal modes, we substitute $u = \bar{u} e^{i\omega t}$, where $\bar{u}$ is a time-independent still undetermined vector, and eventually the actual displacements are identified with the real part of the complex solutions for u. By evaluating the time derivative, we transform the differential equation (5.78) for u into the algebraic equation

$$M_n \omega^2 \bar{u}_\alpha = \sum_\beta \Phi_{\alpha\beta} \bar{u}_\beta \tag{5.79}$$

for $\bar{u}$, which can also be written as

$$\sum_\beta \left(\omega^2 \delta_{\alpha\beta} - \frac{1}{M_n} \Phi_{\alpha\beta} \right) \bar{u}_\beta = 0. \tag{5.80}$$

This is the secular equation for the diagonalization of the matrix $D = M_n^{-1} \Phi$ (the so-called dynamical matrix). The eigenvalues are the values of ω^2 which make the determinant $|\omega^2 \mathscr{I} - D|$ vanish. The eigenvector corresponding to each ω^2 is the nonzero (complex) vector $\bar{u}$ satisfying Eq. (5.80). With its $3N_n$ components, $\bar{u}$ describes the relative amplitude and phase of the displacement of all the coordinates of the N_n atoms. Different masses can be accounted for by a simple generalization of the dynamical matrix.[32]

The present formalism is completely general: it applies equally well to the determination of the normal modes of either a polyatomic molecule (see e.g. Fig. 3.12), or of a crystalline or amorphous solid. In each case, the calculation of the normal frequencies and displacements requires the diagonalization of a $3N_n \times 3N_n$ matrix.

For a crystal, we take advantage of the lattice symmetry to greatly simplify[33] the diagonalization of D. Consider first the case of a single atom per unit cell. We construct a wave-type ansatz for $\bar{u}$: the ξ component of the displacement of atom j (sitting in the cell identified by lattice translation $\mathbf{R}$) is $\bar{u}_{j\xi} = e^{i\mathbf{k}\cdot\mathbf{R}} \varepsilon_\xi$. Here ε_ξ is the $\xi = x$, y, or z component of a still undetermined 3D vector. From Eq. (5.80), we obtain the following equation for ε_ξ:

---

[32] In practice, replacing M_n^{-1} with $M_{nj}^{-1/2} M_{nj'}^{-1/2}$ in the expression for the dynamical matrix D.

[33] This simplification is equivalent to the one which Bloch's theorem provides to the description of the electronic motions in a perfect crystal, Sect. 5.2.

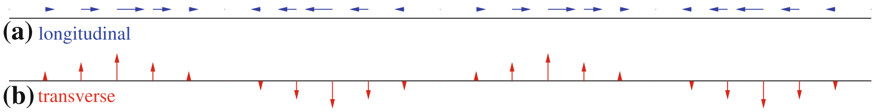


Fig. 5.81 For a given wavelength (i.e. a given **k**), a string exhibits three normal modes of vibration. **a** A longitudinal mode exhibits displacements parallel to **k**. **b** Two transverse modes with displacements perpendicular to **k**; one is drawn, the second one vibrates orthogonally to the plane of the figure

$$\sum_{\xi'} \left[\omega^2 \delta_{\xi\xi'} - \tilde{D}_{\xi\xi'}(\mathbf{k}) \right] \varepsilon_{\xi'} = 0, \tag{5.81}$$

where we exploit the lattice translational invariance of the dynamical matrix

$$D_{j\xi\ j'\xi'} = D_{j''\xi\ j'''\xi'} \quad \text{if} \quad \mathbf{R}_j - \mathbf{R}_{j''} = \mathbf{R}_{j'} - \mathbf{R}_{j'''}, \tag{5.82}$$

to obtain its Fourier transform in reciprocal space:

$$\tilde{D}_{\xi\xi'}(\mathbf{k}) = \sum_j e^{-i(\mathbf{R}_j - \mathbf{R}_{j'})\cdot\mathbf{k}} D_{j\xi\ j'\xi'}, \tag{5.83}$$

which is independent of the choice of $\mathbf{R}_{j'}$. Equation (5.81) shows that the normal modes of a monoatomic crystal are the result of the diagonalization of a 3×3 **k**-dependent matrix. As a consequence, for each **k** point in the first BZ, 3 normal modes and associated frequencies $\omega_s(\mathbf{k})$ (labeled by an index $s = 1, 2, 3$) are obtained. The number of **k** points in the BZ is precisely N_n, see Eq. (5.53): overall we retrieve precisely $3N_n$ normal modes.

The example of the vibrations of a string, drawn in Fig. 5.81, suggests that for a given **k** (here, a given wavelength), three oscillations are observed: one *longitudinal* "compression-dilation" mode associated to displacements parallel to the string, plus two *transverse* modes associated to perpendicular displacements. The same example of the string suggests that for small **k** (long wavelength), the frequencies of all three of these modes should be proportional to $|\mathbf{k}|$. Likewise, for each **k**, the three frequencies of vibration in 3D solids correspond to one (mainly) longitudinal and two (mainly) transverse modes, with an "acoustic" behavior ($\omega \simeq v_\text{s}|\mathbf{k}|$ at small $|\mathbf{k}|$, like sound waves in air). Like for the vibrating string three generally different and $\hat{\mathbf{k}}$-dependent sound velocities v_s, one for each *acoustic branch* are observed. Figure 5.82 sketches this typical dispersion. The zero-frequency acoustic modes at $\mathbf{k} = 0$ are the "infinite-wavelength vibrations" representing the free $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ rigid translations of the solid.

If n_d atoms occupy each crystal unit cell,[34] the problem complicates slightly because of the need to label the displacements of individual atoms within the cell: ξ spans a total $3n_d$ different values, rather than just 3 as for the monoatomic crystal.

---

[34] So that the total number of atoms $N_n = n_d \times N_\text{cell}$, where N_cell is the number of cells in the solid.

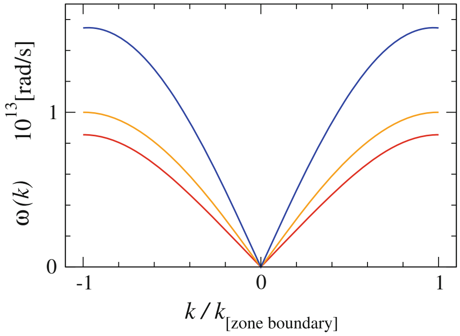


Fig. 5.82 Typical phonon dispersion *curves* along a generic direction in **k**-space for a monoatomic crystal (1-atom basis). The two *lower curves* track the frequencies of the mainly transverse modes, the *upper curve*, the mainly longitudinal mode. The slopes of the three branches for $|\mathbf{k}| \simeq 0$ are the three sound velocities v_s. In some high-symmetry **k**-space direction the two transverse phonon branches could be degenerate

The eigenvector components ε_ξ to be determined are now $3n_d$. These eigenvectors are the result of the diagonalization of a dynamical matrix $\tilde{D}$ of size $3n_d \times 3n_d$ (still enormously smaller than the original $3N_n \times 3N_n$ problem). As a result, for each **k**, we have $3n_d$ modes. Three of these modes are analogous to the acoustic modes of a monoatomic crystal, and represent deformation waves with all atoms in each unit cell oscillating essentially in phase. The remaining $(3n_d - 3)$ modes represent atomic vibrations leaving the cell center of mass immobile. These modes could be visualized by analyzing the simple model problem of a linear chain with springs of alternating strength connecting nearest-neighboring point masses. At all **k**, the dispersion of these $(3n_d - 3)$ *optical modes* remains characteristically far from $\omega = 0$, as illustrated in Fig. 5.83 for a crystal with $n_d = 2$ atoms per cell.

In summary, the harmonic normal modes of a general crystal have frequencies $\omega_s(\mathbf{k})$, and are labeled by **k** plus an index $s = 1, 2, \ldots, 3n_d$ identifying the dispersion

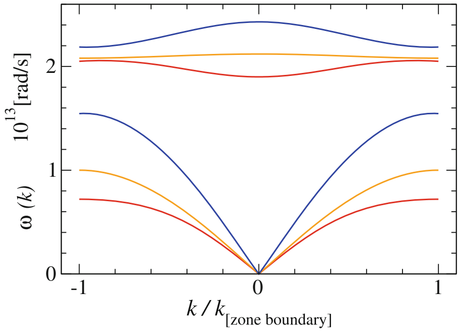


Fig. 5.83 Typical phonon dispersion *curves* along a general direction in **k**-space for a crystal structure with a $n_d = 2$-ions basis. The three *lower curves* (acoustic branches) are approximately linear in $|\mathbf{k}|$ for small $|\mathbf{k}|$. The three *upper curves* (optical branches) become quite flat if intra-cell interactions are much stronger than those between cells

branch. The normal displacements form a basis for the classical dynamics of the whole crystal: arbitrary atomic motions, e.g. wave packets, are linear superpositions of the eigenvectors of $\tilde{D}$ multiplied by the appropriate $e^{i\mathbf{k}\cdot\mathbf{R}}$ factor. The coefficients of the linear combination are fixed by the initial condition.

Now that the classical problem of normal modes is under control, recall that the atomic positions are quantum degrees of freedom rather than classical ones. They should therefore obey Schrödinger equation (3.9) rather than Newton equation (5.77). Promote the normal-mode classical displacements back to quantum operators. Accordingly, replace the $3N_n$ classical oscillations with $3N_n$ quantum harmonic oscillators. A basis state for the Hilbert space of the whole vibrating crystal is then characterized by the number $v_{\mathbf{k}s}$ of excitations (*phonons*) of each oscillator [of frequency $\omega_s(\mathbf{k})$]. In state $|\{v_{\mathbf{k}s}\}\rangle$, the vibrational energy of the whole crystal, viewed as a huge polyatomic molecule, Eq. (4.110), is

$$E_{\text{vib}}(\{v_{\mathbf{k}s}\}) = \sum_{\mathbf{k}s} \hbar\omega_s(\mathbf{k}) \left(v_{\mathbf{k}s} + \frac{1}{2} \right). \tag{5.84}$$

Apart for a less trivial dispersion $\omega_s(\mathbf{k})$ rather than simply $\omega(\mathbf{k}) = c|\mathbf{k}|$, these oscillations resemble closely the oscillations of the electromagnetic fields in a cavity, whose thermodynamics we studied in Sect. 4.3.2.2. This close similarity suggested the name *phonons* for the quanta of these crystal harmonic oscillators. The creation of a phonon of given crystal wave vector $\mathbf{k}$ and branch s simply means raising the quantum number $v_{\mathbf{k}s}$ of the oscillator $(\mathbf{k}, s)$ by one, thus increasing the total vibrational energy by $\hbar\omega_s(\mathbf{k})$.

Phonon dispersions are investigated by means of several techniques, including *neutron inelastic scattering*, ultrasound echo, Raman, and IR absorption. All these techniques involve excitations or deexcitation of one or several phonon states. The outcome of these experiments generally agrees with the sophisticated calculations based on the evaluation and diagonalization of the dynamical matrix D. Figures 5.84 and 5.85 show examples of measured phonon dispersions of real materials. Optical phonons often appear as absorption peaks in the IR spectrum of otherwise transparent insulators. This absorption is especially strong for ionic crystals, e.g. NaCl or LiF (Fig. 5.86), where optical displacements are associated to a large electric-dipole change.

5.3.2 Thermal Properties of Phonons

Phonons are essential to understand the thermal properties of solids. In particular, the heat capacity of all solids is mainly associated to the crystal vibrations.[35] By further developing the analogy of photons and phonons, i.e. by applying the statistics of

---

[35] With a few noteworthy exceptions: (i) in metallic solids, at very low temperature (typically few K) the tiny electronic contribution (Fig. 4.15) dominates over the phonon heat capacity;

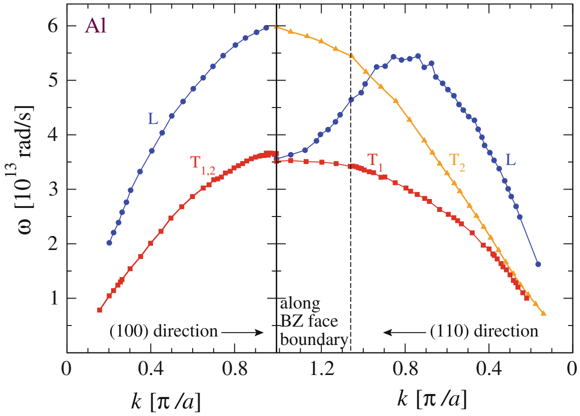


Fig. 5.84 Inelastic neutron scattering measurements of the phonon dispersion of crystalline (fcc) Al, along two k-space lines: (100) and (110). Note that the two transverse branches T_1 and T_2 remain degenerate along the (100) line, but exhibit distinct frequencies along the (110)

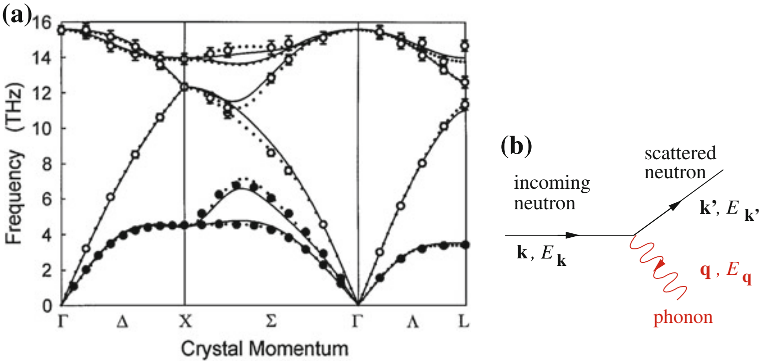


Fig. 5.85 **a** The observed phonon dispersion of silicon, measured by inelastic neutron (*circles*) and by X-ray (*solid line*) scattering. Note the presence of both acoustical and optical branches, as appropriate for a (diamond) structure with a basis of $n_d = 2$ atoms per cell. The letters mark special points in **k**-space, e.g. Γ is $\mathbf{k} = \mathbf{0}$. Along high-symmetry paths (Γ-X and Γ-L) pairs of transverse branches are degenerate, while this degeneracy is resolved along other directions. Reprinted figure with permission from M. Holt, Z. Wu, H. Hong, P. Zschack, P. Jemian, J. Tischler, H. Chen, and T.-C. Chiang, Phys. Rev. Lett. **83**, 3317 (1999). Copyright (1999) by the American Physical Society. **b** The scheme of the inelastic scattering event

the harmonic oscillator as in Eq. (4.62), the total internal energy of the harmonic crystal is

---

(Footnote 35 continued)
(ii) occasionally, localized magnetic moments add a contribution which may become important at low temperature.

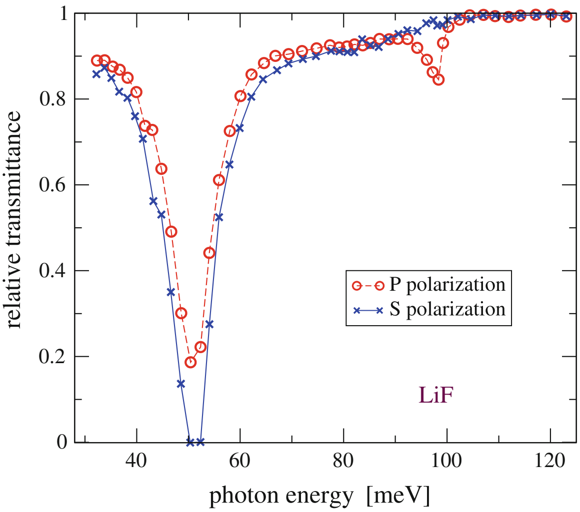


Fig. 5.86 Observed transmittance of IR radiation by a 200 nm-thick LiF film at room temperature. Radiation impinges on the film at ~30° from vertical. The electric field is parallel/perpendicular to the plane of incidence for of P/S-polarized radiation, respectively. The peaks measure the frequencies of the transverse and longitudinal optical phonons, at energy $\hbar\omega \simeq 50\,\text{meV}$ and $\hbar\omega \simeq 100\,\text{meV}$ respectively. Momentum conservation [Eq. (5.74) and Fig. 5.85b] implies that IR radiation absorption probes the $\mathbf{k} \simeq \mathbf{0}$ region of the optical phonon branches (Data from Ref. [47])

$$U_{\text{vib}} = U_0 + \sum_{\mathbf{k}\,s} \varepsilon_{\mathbf{k}\,s} [n_{\varepsilon_{\mathbf{k}s}}]_B = U_0 + \sum_{\mathbf{k}s} \frac{\varepsilon_{\mathbf{k}s}}{\exp\left(\frac{\varepsilon_{\mathbf{k}s}}{k_B T}\right) - 1}, \quad \text{with} \quad \varepsilon_{\mathbf{k}s} = \hbar\omega_s(\mathbf{k}). \tag{5.85}$$

U_0 includes the zero-point energy $\sum_{\mathbf{k}s} \hbar\omega_s(\mathbf{k})/2$, which is a material-dependent constant, irrelevant for thermodynamics. Using for each oscillator the result of Eq. (4.63), the T-derivative of U_{vib} yields the heat capacity:

$$C_{V\,\text{vib}} = k_B \sum_{\mathbf{k}s} \frac{x_{\mathbf{k}s}^2 \exp(x_{\mathbf{k}s})}{\exp(x_{\mathbf{k}s}) - 1} = k_B \sum_{\mathbf{k}s} \left[\frac{x_{\mathbf{k}s}/2}{\sinh(x_{\mathbf{k}s}/2)}\right]^2, \quad \text{with} \quad x_{\mathbf{k}s} = \frac{\hbar\omega_s(\mathbf{k})}{k_B T}. \tag{5.86}$$

Equations (5.85) and (5.86) depend on the detail of the phonon dispersions. In practice however, experimentally crystals exhibit a fairly universal T-dependence of the molar specific heat (once the T-linear electronic contributions of metals is subtracted) illustrated in Fig. 5.87: a $\propto T^3$ raise at low temperature (see also Fig. 4.15), and stabilization to a constant $\simeq 3N_A k_B$ (Dulong-Petit) limit at high temperature.

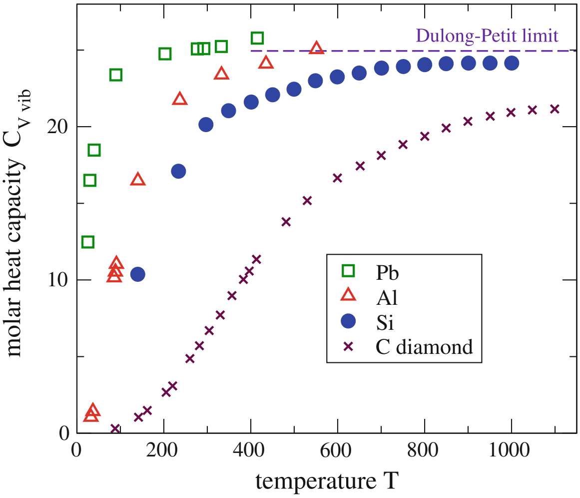


Fig. 5.87 The measured molar heat capacity at constant volume for a few solids. At high temperature (but limited by melting), $C_{V\,\mathrm{vib}}$ approaches the universal $3N_Ak_B \simeq 24.9\,\mathrm{J\,mol^{-1}\,K^{-1}}$ Dulong-Petit value

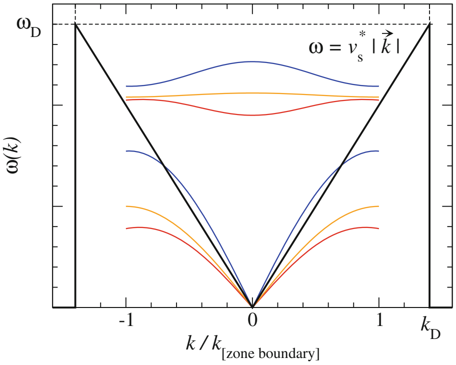


Fig. 5.88 The Debye model approximates the phonon branches of a real solid with an effective linear dispersion $\omega \simeq v_s^* |\mathbf{k}|$. The sound velocity v_s^* is taken as a suitable average of the $\mathbf{k} \to 0$ slopes of the three acoustic branches over all directions. The real phonon modes is replaced by integration over a sphere of radius k_D and multiplication by a factor 3 representing the three acoustic branches, which are assumed degenerate. The wave-number cutoff k_D is of the same order as the size of the BZ, but it becomes significantly larger for a crystal structure with several atoms per cell, to account for multiple optical branches

To capture this T dependency, a popular simplified model is usually adopted: the *Debye model*. The details of the phonon dispersions are replaced by a fictitious linear dispersion (Fig. 5.88) reproducing the minimal quantitative properties of the phonons of the material under consideration:

1. The average speed of sound v_{s}^{*}. This average slope of the acoustic branches close to $\mathbf{k} = \mathbf{0}$ is chosen such as to provide the correct factor Ξ in the low-energy density of oscillator energies $g(\varepsilon) = \Xi\, \varepsilon^2$.[36] In this way the model fits the low-T limit, where only low-energy oscillators are excited significantly.
2. The total number of vibrational modes must equal $3N_n$, in order to guarantee the correct high-temperature behavior, in accord with classical equipartition.

The latter point is implemented by replacing the sum over $\mathbf{k}$ and s in (5.85) by a sum over $\mathbf{k}$ extended over a spherical region of radius k_{D} (the Debye wave number) and a factor $g_s = 3$ representative of the three acoustic branches.[37] The allowed $\mathbf{k}$ points of a finite sample of N_n atoms in a $L \times L \times L$ cube are $\mathbf{k} = \frac{2\pi}{L}\,\mathbf{n}$ [see Eqs. (4.38) and (5.53)]. This observation allows us to determine the radius k_{D} of the sphere of allowed oscillators, as follows:

$$3N_n = g_s \sum_{|\mathbf{n}|<n_{\mathrm{D}}} 1 \simeq g_s \int_0^{n_{\mathrm{D}}} n^2 dn = g_s \frac{4\pi}{3} n_{\mathrm{D}}^3, \tag{5.87}$$

whence $n_{\mathrm{D}} = \left(\frac{3}{4\pi} N_n\right)^{1/3}$, and thus $k_{\mathrm{D}} = \frac{2\pi}{L} n_{\mathrm{D}} = \left(6\pi^2 \frac{N_n}{V}\right)^{1/3}$. This corresponds to a Debye cutoff energy

$$\varepsilon_{\mathrm{D}} = \hbar\omega_{\mathrm{D}} = \hbar v_{\mathrm{s}}^{*} k_{\mathrm{D}} = \hbar\, v_{\mathrm{s}}^{*} \left(6\pi^2 \frac{N_n}{V}\right)^{1/3}. \tag{5.88}$$

As established in Eq. (4.114), the energy-density of states corresponding to a dispersion $\omega = v_{\mathrm{s}}^{*}|\mathbf{k}|$ is quadratic in ε:

$$g_s g_{\mathrm{ph}}(\varepsilon) = \begin{cases} 9N_n\, \varepsilon^2/\varepsilon_{\mathrm{D}}^3 & \text{for} \quad 0 \le \varepsilon \le \varepsilon_{\mathrm{D}} \\ 0 & \text{elsewhere} \end{cases}. \tag{5.89}$$

With this approximate (Debye) density of oscillators, the internal energy Eq. (5.85) becomes

$$U_{\mathrm{vib}} = U_0 + \int_0^{\varepsilon_{\mathrm{D}}} g_s g_{\mathrm{ph}}(\varepsilon) \frac{\varepsilon}{e^{\frac{\varepsilon}{k_{\mathrm{B}}T}} - 1} d\varepsilon = U_0 + \frac{9N_n}{\varepsilon_{\mathrm{D}}^3} \int_0^{\varepsilon_{\mathrm{D}}} \frac{\varepsilon^3}{e^{\frac{\varepsilon}{k_{\mathrm{B}}T}} - 1} d\varepsilon. \tag{5.90}$$

As $9N_n/\varepsilon_{\mathrm{D}}^3 = {}^3\!/_2 V/(\pi^2\hbar^3 v_{\mathrm{s}}^{*3})$, Eq. (5.90) reproduces formally the expression (4.115) for photons, except for (i) a factor $g_s = 3$ rather than 2, (ii) the sound velocity

---

[36] In view of the speed dependence in Eq. (4.114), v_{s}^{*} is obtained as the angular average of the inverse cube of the three speeds of sound: $3v_{\mathrm{s}}^{*-3} = \int (v_{\mathrm{s}1}^{-3} + v_{\mathrm{s}2}^{-3} + v_{\mathrm{s}3}^{-3}) d\Omega_{\hat{\mathbf{k}}}/(4\pi)$.

[37] The eventuality of optical branches is effectively accounted for by the value of the cutoff k_{D}, which is fixed precisely by imposing that the total number of modes is $3N_n$.

v_s^* in place of c, and (iii) the cutoff ε_D in the energy integration. Equation (5.90) is rewritten conveniently as

$$U_{\text{vib}} = U_0 + 9N_n \frac{(k_B T)^4}{\varepsilon_D^3} \int_0^{\frac{\Theta_D}{T}} \frac{x^3}{e^x - 1} dx = U_0 + 9N_n k_B \frac{T^4}{\Theta_D^3} \int_0^{\frac{\Theta_D}{T}} \frac{x^3}{e^x - 1} dx, \tag{5.91}$$

where, unsurprisingly, $\Theta_D = \varepsilon_D / k_B$.

At low temperature $T \ll \Theta_D$, the integration end Θ_D/T can be approximated with infinity, and the resulting energy integral over the Planck distribution is the same as in Eq. (4.115): $\int_0^\infty x^3/(e^x - 1)dx = \pi^4/15$. In this limit, the internal energy is therefore

$$U_{\text{vib}}(T \ll \Theta_D) \simeq U_0 + 9N_n k_B \frac{T^4}{\Theta_D^3} \frac{\pi^4}{15} = U_0 + \frac{3\pi^4}{5} N_n k_B \frac{T^4}{\Theta_D^3}. \tag{5.92}$$

Derivation with respect to T yields a heat capacity

$$C_{V\,\text{vib}}(T \ll \Theta_D) \simeq \frac{12\pi^4}{5} N_n k_B \left(\frac{T}{\Theta_D} \right)^3, \tag{5.93}$$

reproducing the experimental behavior at low temperature.

For large temperature, the ratio Θ_D/T is small, and so is the variable x in Eq. (5.91). The integrand function is then expanded as $x^3/(e^x - 1) \simeq x^2$. In this limit, the integral in Eq. (5.91) yields approximately $\frac{1}{3}\,(\Theta_D/T)^3$, thus the internal energy is

$$U_{\text{vib}}(T \gg \Theta_D) \simeq U_0 + 9N_n k_B \frac{T^4}{\Theta_D^3} \frac{1}{3} \left(\frac{\Theta_D}{T} \right)^3 = U_0 + 3N_n k_B T. \tag{5.94}$$

This form matches the classical equipartition limit, corresponding to a heat capacity

$$C_{V\,\text{vib}}(T \gg \Theta_D) \simeq 3N_n k_B, \tag{5.95}$$

reproducing the Dulong-Petit limit.

The general expression, accurate at arbitrary temperature, derives directly from Eqs. (5.86) and (5.89):

$$C_{V\,\text{vib}}(T) = 9N_n k_B \left(\frac{T}{\Theta_D} \right)^3 \int_0^{\frac{\Theta_D}{T}} \left[\frac{x^2/2}{\sinh(x/2)} \right]^2 dx. \tag{5.96}$$

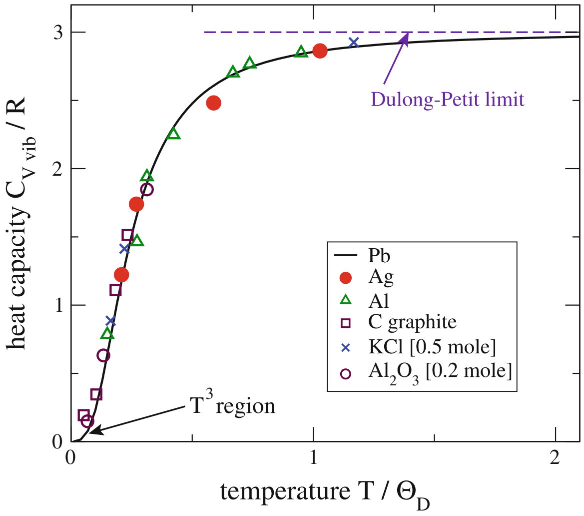


Fig. 5.89 The measured heat capacity (in units of $R = k_B N_A$) of five individual solids. Temperature is rescaled by the Debye temperature Θ_D of each material, so that all data follow the universal (Debye) *curve* of Eq. (5.96) (*solid line*). For a better comparison, here $C_{V\text{vib}}$ data refer to $N_n = N_A$ atoms in total: this represents 1 mole of Ag, Al, and C, but 0.5 mole of KCl (2 atoms per chemical formula), and 0.2 mole of Al_2O_3 (5 atoms per formula) (Inspired by Fig. 13-9 of Ref. [3])

The Debye form (5.96) interpolates between the low- and high-temperature limits, see Fig. 5.89. The T dependence of $C_{V\text{vib}}(T)$ agrees closely with most experimental data of crystalline solids. This agreement on *thermodynamics* is striking, for a model which gathers all the complications of the phonon dispersions in a single parameter Θ_D (or, equivalently, ε_D, or ω_D) that can be determined purely through the measurement of *mechanical* properties of the solid, see Eq. (5.88).

The success of the Debye model yields a central role to its one parameter, as the energy scale characterizing the phonons of a crystal. Table 5.3 lists the Debye temperature of a few examples of crystals. Values of Θ_D range in the same scale as the typical vibrational temperature Θ_{vib} of diatomic molecules (Sect. 4.3.1.2).

The Debye model exhibits the same general "quantum-unfreezing" behavior as a single oscillator (Fig. 4.8). The main difference, the low-temperature $C_{V\text{vib}} \propto T^3$ dependence in the solid [rather than $C_{V\text{vib}} \propto \exp(-\Theta/T)$], is due to the presence of

Table 5.3 The Debye temperature of a few elemental solids

Element	Θ_D (K)	Element	Θ_D (K)
Ag	225	Ge	366
Au	165	Na	159
C (diamond)	1,860	Ni	456
Cu	339	Pt	229

Table 5.4 Compared analytic expressions for the low- and high-temperature heat capacity of several thermodynamic systems relevant for the physics of matter

Thermodynamic system	Low temperature	High temperature
Ideal spin-s system	$N k_{\mathrm{B}} \left(\frac{g\mu_{\mathrm{B}} B}{k_{\mathrm{B}} T}\right)^2 \exp\left(-\frac{g\mu_{\mathrm{B}} B}{k_{\mathrm{B}} T}\right)$	$\frac{1}{3} s(s+1) N k_{\mathrm{B}} \left(\frac{g\mu_{\mathrm{B}} B}{k_{\mathrm{B}} T}\right)^2$
Rotations of molecules	$3N k_{\mathrm{B}} \left(\frac{\hbar^2}{I k_{\mathrm{B}} T}\right)^2 \exp\left(-\frac{\hbar^2}{I k_{\mathrm{B}} T}\right)$	$N k_{\mathrm{B}}$
Vibrations of molecules	$N k_{\mathrm{B}} \left(\frac{\hbar\omega}{k_{\mathrm{B}} T}\right)^2 \exp\left(-\frac{\hbar\omega}{k_{\mathrm{B}} T}\right)$	$N k_{\mathrm{B}}$
Vibrations of crystals	$\frac{12\pi^4}{5} N k_{\mathrm{B}} \left(\frac{T}{\Theta_{\mathrm{D}}}\right)^3$	$3N k_{\mathrm{B}}$
Electromagnetic fields	$\frac{4\pi^2}{15} V k_{\mathrm{B}} \left(\frac{k_{\mathrm{B}} T}{\hbar c}\right)^3$	(for any T)
Gas of ideal fermions	$\frac{\pi^2}{2} N k_{\mathrm{B}} \frac{T}{T_{\mathrm{F}}}$	$\frac{3}{2} N k_{\mathrm{B}}$

a continuum of oscillators of arbitrarily low frequency: even at low temperature, all oscillators with $\hbar\omega < k_{\mathrm{B}} T$ remain quantum-unfrozen and contribute k_{B} to $C_{V\,\mathrm{vib}}$.

It is instructive to compare the temperature dependence of the heat capacity of the Debye model to that of other thermodynamic systems studied earlier on. In particular, the low- and high-T temperature dependences collected in Table 5.4 deserve a thorough understanding.

5.3.3 *Other Phonon Effects*

In the absence of electronic mechanisms, phonons are the responsible of heat transport in insulating solids. With arguments of the type employed for electronic heat transport in Sect. 5.2.2.1, it is possible to evaluate heat conduction by phonons, and to account quantitatively for the observational data [10, 11].

Macroscopic elastic properties of solids are directly related on the small-**k** region of the acoustic phonon branches, i.e. on the sound velocities. Small sound velocities call for soft easily-deformable solids; large sound velocity are typical of rigid materials.

Thermal expansion is also associated to atomic displacements. Harmonic phonons however do not account for any expansion, because $\langle v_{\mathbf{k}s} | u_{\mathbf{k}s} | v_{\mathbf{k}s} \rangle = 0$ for any phonon number $v_{\mathbf{k}s}$, thus at any temperature $[u_{\mathbf{k}s}] = 0$: a temperature raise only increases the average amplitude of oscillation around the same fixed equilibrium separation. Thermal expansion arises therefore entirely as a consequence of the anharmonicity of the potential, i.e. of terms $\mathscr{O}(u^3)$ and higher-order in Eq. (5.75), which are present in real solids but are neglected in the harmonic theory of phonons.

Finally, displacements of the atoms away from their ideal crystal positions associated to phonons affect structural determinations done by means of X-ray and neutron diffraction experiments. For not too low T, the amplitude of the random thermal atomic displacements $[u_\alpha^2] \propto T$. These random movements leave the diffracted peak positions unchanged, but remove some intensity from the peaks, and transfer it into a diffuse background.

Problems

A ⋆ marks advanced problems.

5.1 In the approximation of free non-interacting electrons at temperature 0 K compute the mean speed $\langle |v| \rangle$ of a conduction electron in solid aluminum, using the fact that X-rays of wavelength $\lambda = 100\,\text{pm}$ are diffracted by its fcc structure at a minimum angle $2\theta = 24.7°$, and recalling that every Al atom contributes 3 electrons to the conduction band.

5.2 Solid titanium has a density 4,507 kg/m^3. Sound waves propagate through titanium at an average speed $v_s = 4.14\,\text{km/s}$. In the Debye approximation, determine the low-temperature molar heat capacity of the phonons of titanium, specifically evaluating the parameters α and A which express its temperature dependence in the relation $C_v = AT^\alpha$. Evaluate also the molar heat capacity at a such a high temperature that the classical limit provides a good approximation for C_v. Finally, determine at what temperature the heat capacity reaches 1 % of its maximum value.
[Recall the relation connecting the Debye cutoff frequency to the average speed of sound v_s and the number density N_at/V of the atoms in the solid: $\omega_D = v_s \left(6\pi^2 N_\text{at}/V\right)^{1/3}$.]

5.3 For the energy of the conduction band of a hypothetical one-dimensional cesium crystal, take the expression $\varepsilon_k = -A - B\cos(ka)$, where $A = 2.9\,\text{eV}$, $B = 4.1\,\text{eV}$, k is the wave vector, and $a = 125\,\text{pm}$. The expression for ε_k is valid in an energy scale whose zero refers to an electron at rest far away from solid. Determine the maximum wavelength of electromagnetic radiation capable to produce electron photoemission from such a solid.

5.4 ⋆ Consider a one-dimensional crystal of lattice spacing a. Estimate the lowest-energy band gap at the BZ boundary in the approximation of quasi-free electrons (plane-waves basis), and assuming that the effective periodic potential acting on the electron takes the form

$$V(x) = E\left(\cos\frac{2\pi x}{a} + \frac{1}{3}\cos\frac{6\pi x}{a}\right),$$

where $E = 6.8\,\text{eV}$. If each atom contributes two electrons, is the crystal a metal or an insulator? Is it opaque or transparent for visible light?

5.5 ⋆ Three different crystalline solids A, B and C have such structures that in each of them 2 different atoms (total mass $M_1 + M_2 = 120$ atomic mass units) occupy every primitive cell of the respective periodic lattice. A has simple-cubic (sc) structure; B has face-centered cubic (fcc) structure; C has body-centered cubic (bcc) structure. Neutrons of 10 meV kinetic energy diffract over powder samples of A, B, and C, forming diffraction rings at several angles. For all three samples the smallest diffracted angle is $2\theta = 32.4°$. Evaluate the side a of the conventional cubic cells for each one of the 3 solids, determine which solid has the largest density, and the value of this density.

5.6 The vibrational molar heat capacity C_v of aluminum is 0.0239 J/(mol K) and 0.0804 J/(mol K) at temperatures 20 and 30 K respectively. Knowing that X-rays of wavelength $\lambda = 110$ pm are diffracted by its fcc structure at a minimum angle $2\theta = 18.1°$, evaluate the mean speed of sound v_s in this solid.

[Recall the relation connecting the Debye cutoff frequency to the speed of sound v_s and the number density N_{at}/V of the atoms in the solid: $\omega_D = v_s \left(6\pi^2 N_{\mathrm{at}}/V\right)^{1/3}$.]

5.7 ⋆ Consider a bcc iron crystal. Focus on the phonons characterized by rigid displacements of each atomic plane perpendicular to one of the main diagonals of the conventional cubic cell. Restrict further to the displacements u_i perpendicular to the planes themselves, i.e. longitudinal phonons in the (111) direction. Assume an elastic interaction between near-neighbor planes, i.e. a restoring force acting on each atom in the ith plane given by $F_i = C(u_{i+1} - u_i) + C(u_{i-1} - u_i)$, and indicate with a the equilibrium distance between near-neighbor planes. Write the dispersion relation of such phonons, and evaluate their maximum frequency, in the hypothesis that $C = 65.6$ N/m.

5.8 The phenomenon of Bloch oscillations expected in an ideal crystal could be observed even in a real crystal, provided that the mean collision time τ of the conduction electrons is long enough. Assuming that a uniform constant electric field $|\mathbf{E}| = 1$ V/m is applied to a sodium metal sample, estimate the minimum value of τ for an electron to stand a chance to cross the entire BZ without colliding. Assume for simplicity a one-dimensional monoatomic crystal of lattice spacing 200 pm.

5.9 Compute the smallest nonzero Bragg diffraction angle θ observed in neutron elastic-scattering experiments. The neutron kinetic energy is 22 meV. The sample is microcrystalline aluminum of face-centered cubic (fcc) structure, density 2,700 kg m^{-3}, and atomic mass number $A = 27$.

5.10 The following molar heat capacity of gallium arsenide ($^{70}\mathrm{Ga}^{75}\mathrm{As}$, density $\rho = 5{,}320\,\mathrm{kg/m^3}$) were measured:

T (K)	C_V ($\mathrm{J\,K^{-1}\,mol^{-1}}$)
1.0	8.33×10^{-5}
2.0	6.67×10^{-4}
3.0	2.25×10^{-3}

Based on these data, evaluate the average sound velocity according to the Debye model and extrapolate the molar heat capacity at $T = 1{,}000\,\mathrm{K}$.
[Recall the Debye-model expression for the vibrational internal energy *per atom* in the Debye model:

$$U_1 = 9k_\mathrm{B}\frac{T^4}{\Theta_D^3}\int_0^{\Theta_D/T}\frac{x^3dx}{e^x-1} \quad \text{and the result} \quad \int_0^{\infty}\frac{x^3dx}{e^x-1} = \frac{\pi^4}{15}.]$$

Appendix A
Conclusions and Outlook

The adiabatic approximation sketched in Sect. 3.1 is a central tool for understanding the combined dynamics of electrons and nuclei. It works well in such contexts, as close-shell molecules and wide-gap insulators, where an excitation gap "protects" the electronic ground state against transitions induced by atomic motion.

Metals violate this adiabatic paradigm since both electrons and vibrations involve arbitrarily small excitation energies. For this reason, even at low temperature, each electron near the Fermi surface has a significant chance of colliding with some phonon (Sect. 5.2.2.1): as a result, in contrast to the adiabatic assumption, the overall electronic state of a metal changes all the time. Precisely the absence of a gap in the electronic spectrum makes the neglect of nonadiabatic terms of Eq. (3.6) not particularly well justified: the adiabatic approximation seems therefore rather ill-grounded in metals. In fact, despite these difficulties, it still makes sense to apply the adiabatic concepts to metals, since phonon scattering involves only a small fraction of the electrons: those within meVs of the chemical potential. All other "deeper" electrons remain "frozen" in their adiabatic state due to the forbidden (by Pauli's principle) "sea" of already occupied states sitting at near energies. This protection is equivalent to an effective gap, of the order of the energy distance of these deeper electrons to the chemical potential. This vast majority of (deeper) electrons is the main responsible of the crystal cohesion, and thus of the existence of a fairly well-defined adiabatic potential acting on the ions. It is precisely this adiabatic potential that provides a stable crystal structure and phonons even for a metal. The frequent collisions of Fermi-surface electrons with the vibrating crystal are then described as *electron-phonon coupling*, associated to the nonadiabatic terms, and result in a relatively weak perturbation to the adiabatic motion. Eventually however, the ground state of over one third of the elemental metals, and of many metallic compounds cannot be simply interpreted as an adiabatically separated state, but rather a *superconducting state*. While the material remains a solid (due to the cohesion produced by its deeper electrons), in a superconductor a correlated motion of the phonons and the Fermi-surface electrons yields peculiar collective thermodynamic and transport properties, quite distinct from those of the ordinary resistive metallic state.

N. Manini, *Introduction to the Physics of Matter*,
Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-14382-8

Furthermore, in many materials, electron-electron correlations introduce new physics beyond the independent-electron approximation. Magnetism, basically a consequence of the same repulsion-induced correlations as Hund's rules (Sect. 2.2.8.3) in atoms, is a macroscopic deviation from the band picture, where, strictly speaking, it finds no place (see Fig. 5.53).

The present introductory course leaves out a number of important topics, including the coexistence of different phases, order parameters, phase transitions, chemical reactions, amorphous and nanostructured solids, the liquid, superfluid, and plasma states of matter, and all sort of *collective behavior* which makes the study of the thermodynamics of matter fascinating. The account of experimental techniques is also radically limited to a few providing the most fundamental and transparent datasets. Readers are likely to encounter several other popular techniques in lab courses. A solid understanding of the material in the present textbook should provide the reader with the conceptual tools for understanding the investigated physical phenomena, and estimating the resolution and the limits of several specific classes of experimental techniques.

Appendix B
Elements of Quantum Mechanics

We summarize here the main concepts and ingredients of QM needed to digest the physics of matter in the present volume. The present level of treatment is quite basic and compact, aiming at providing the minimal conceptual tools needed to understand the physics of matter. Most subtle epistemologic implications[1] and mathematical derivations are omitted. This Appendix collects and organizes an absolute minimum of QM concepts. A student in physics will probably enjoy the present textbook better after taking a full semester course in QM, or reading substantial parts e.g. of one of these textbooks: Refs. [5, 15–17].

B.1 Bras, Kets and Probability

In QM, a physical state is represented by a unit-norm vector—a *ket* $|j\rangle$ in Dirac notation [17]—in an abstract complex Hilbert space of vectors. The symbol j identifying uniquely the quantum state stands for a full set of *quantum numbers*. A *bra* $\langle q'|$ is the linear operator which, given any state $|j\rangle$, yields its projection (or "overlap") $\langle q'|j\rangle$ along the "direction" representing state $|q'\rangle$. This projection is a complex number. The bra-ket application defines an inner product between vectors in the Hilbert space. The special case $\langle j|j\rangle$ of the overlap of a vector with itself produces a real nonnegative number, which represents the squared norm of the vector itself. When a vector represents a physical state, its norm equals 1. As long as $|j\rangle$ is not the null vector, it can be normalized by dividing it by its norm: $|j\rangle/(\langle j|j\rangle)^{1/2}$. The basic property of kets in a Hilbert space is linearity: $a|j\rangle + a'|j'\rangle$ is also a ket, for arbitrary complex

---

[1] QM can be far from intuitive. Two great gurus of QM used to say: *For those who are not shocked when they first come across quantum theory cannot possibly have understood it.* [Niels Bohr, quoted in Heisenberg, Werner (1971). Physics and Beyond. New York: Harper and Row. pp. 206.]; *I think I can safely say that nobody understands quantum mechanics.* [Richard Feynman, in The Character of Physical Law (1965)].

N. Manini, *Introduction to the Physics of Matter*,
Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-14382-8

numbers a and a'. The transformation of a ket into a bra is antilinear, in the sense that the bra associated to $a|j\rangle + a'|j'\rangle$ is $a^*\langle j| + a'^*\langle j'|$, where a^* indicates the complex conjugate to a.

Any observable, i.e. any quantity which can be measured in a physical experiment, is associated with a self-adjoint linear operator Q from the Hilbert space into itself. The fact that Q brings the Hilbert space into itself means that for any ket $|j\rangle$, $Q|j\rangle$ is a (generally unnormalized) ket too. The *linearity* of Q means simply that $Q(a|j\rangle + a'|j'\rangle) = aQ|j\rangle + a'Q|j'\rangle$ for any complex numbers and kets involved. The adjoint $X^\dagger$ of a linear operator X is a generalization of the complex conjugation which comes about when converting $X|j\rangle$ to the corresponding bra: $X|j\rangle \rightarrow \langle j|X^\dagger$. A *self-adjoint* operator Q is such that $Q^\dagger = Q$. In practice the self-adjointness means that $\langle j|Q|j'\rangle = \langle j'|Q|j\rangle^*$ for any $|j\rangle$ and $|j'\rangle$.

The values which may come up as the result of a single measurement of the observable associated to Q are the *eigenvalues* of the operator Q, namely those special numbers q' such that $Q|q'\rangle = q'|q'\rangle$. Due to the self-adjointness property of Q, its eigenvalues are *real* quantities. Moreover, the set $\{|q'\rangle\}$ of all possible *eigenstates* of Q, is a complete set of orthogonal states. This means that any state in the Hilbert space can be written as a linear combination $|j\rangle = \sum_{q'} a_{q'}|q'\rangle$. Assuming that all $|q'\rangle$ are properly normalized ($\langle q'|q'\rangle = 1$), the coefficients in the linear combination are $a_{q'} = \langle q'|j\rangle$.

These coefficients bear an important physical significance. The probability that, with the system initially in state $|j\rangle$, a measurement of Q yields the eigenvalue q is the square module of the overlap of the physical state $|j\rangle$ with the relevant eigenket of Q: $P(q) = |\langle q|j\rangle|^2$. In case q is degenerate, i.e. the same eigenvalue is associated to several orthonormal eigenkets $|q_i\rangle$, then one needs to sum over these eigenkets:

$$P(q) = \sum_{\substack{i \\ q_i = q}} |\langle q_i|j\rangle|^2 . \tag{B.1}$$

The probability distribution $P(q)$ needs to be properly normalized. Indeed, the sum of $P(q)$ over all possible outcomes q of a measurement must be

$$1 = \sum_q P_j(q) = \sum_{q'} |\langle q'|j\rangle|^2 = \sum_{q'} \langle j|q'\rangle\langle q'|j\rangle = \langle j| \left[\sum_{q'} |q'\rangle\langle q'| \right] |j\rangle . \tag{B.2}$$

The quantity in square brackets is an operator from the Hilbert space into itself, and the completeness of the basis $\{|q'\rangle\}$ guarantees that it coincides with the identity $\mathscr{I}$. This leaves $\langle j|j\rangle$ which must then be unity for the probability to be properly normalized.

B.1.1 Wavefunction and Averages

We obtain an important example of this sort of overlaps when for Q we select the position operator R, with eigenvalues r' and eigenkets $|r'\rangle$. The overlap $\psi_j(r') = \langle r'|j\rangle$ is called the *wavefunction* of state $|j\rangle$ at point r'. $\psi_j(r')$ is a complex number, such that $|\psi_j(r')|^2$ is the probability density $P_j(r')$ that the particle in an initial state $|j\rangle$ is found at r' when its position is measured. Since the eigenvalues of the position operator are continuous quantities, the sum in Eq. (B.2) is replaced by an integral. Thus the normalization of the probability distribution in r'-space is

$$1 = \int P_j(r')\, dr' = \int |\langle r'|j\rangle|^2\, dr' = \int |\psi_j(r')|^2\, dr' \,. \tag{B.3}$$

Going back to a general observable Q, typically one can prepare the quantum system every time in the same initial state $|j\rangle$, repeatedly measure the observable Q, and average over these measurements. Given the individual probabilities for all possible outcomes q' of the measurement of Q, quantum theory provides a straightforward formula to predict this average value $\langle Q\rangle$:

$$\begin{aligned}\langle Q\rangle = \sum_{q'} q' P_j(q') &= \sum_{q'} q' |\langle q'|j\rangle|^2 = \sum_{q'} \langle j|q'\rangle q' \langle q'|j\rangle \\ &= \sum_{q'} \langle j|q'\rangle \langle q'|q'|q'\rangle \langle q'|j\rangle \,. \end{aligned} \tag{B.4}$$

Since $q'|q'\rangle = Q|q'\rangle$,

$$\langle Q\rangle = \sum_{q'} \langle j|q'\rangle \langle q'|Q|q'\rangle \langle q'|j\rangle = \sum_{q} \sum_{q'} \langle j|q\rangle \langle q|Q|q'\rangle \langle q'|j\rangle \,. \tag{B.5}$$

In the last passage we have added a lot of null terms, namely all those with $q \neq q'$. Now, we notice that we have expressed the identity at both sides of Q:

$$\langle Q\rangle = \langle j| \left[\sum_{q} |q\rangle\langle q| \right] Q \left[\sum_{q'} |q'\rangle\langle q'| \right] |j\rangle = \langle j|Q|j\rangle \,. \tag{B.6}$$

We conclude that the average of an observable Q is provided by the diagonal matrix element of its operator on the probed state. Notice that $\langle Q\rangle$ need not coincide with any of the eigenvalues q'! For example, if the outcome of an observable is $+1/2$ one half of the times and $-1/2$ one half of the times, its average equals zero. However zero can occur in no measurement.

B.2 Position, Momentum, and Translations

QM is about the motion of particles. The simplest way to move a particle is to translate it. Consider an eigenstate of position $R|\mathbf{r}'\rangle = \mathbf{r}'|\mathbf{r}'\rangle$, where $\mathbf{r}' = (x', y', z')$. A translation by δx in the $\hat{\mathbf{x}}$ direction modifies $|x', y', z'\rangle$ into $|x' + \delta x, y', z'\rangle$. Indicate with $\mathscr{T}(\boldsymbol{\delta}\mathbf{x})$ the operator for an infinitesimal translation by $\boldsymbol{\delta}\mathbf{x} = \hat{\mathbf{x}}\,\delta x$:

$$|\mathbf{r}'\rangle \;\to\; \mathscr{T}(\boldsymbol{\delta}\mathbf{x})|\mathbf{r}'\rangle = |\mathbf{r}' + \boldsymbol{\delta}\mathbf{x}\rangle = |x' + \delta x, y', z'\rangle\,. \tag{B.7}$$

This result can be generalized to an arbitrary state $|j\rangle$, by expanding it on the basis of the position eigenstates:

$$|j\rangle \;\to\; \mathscr{T}(\boldsymbol{\delta}\mathbf{x})|j\rangle = \mathscr{T}(\boldsymbol{\delta}\mathbf{x}) \int d^3r'|\mathbf{r}'\rangle\langle\mathbf{r}'|j\rangle = \int d^3r'[\mathscr{T}(\boldsymbol{\delta}\mathbf{x})|\mathbf{r}'\rangle]\psi_j(\mathbf{r}')\,. \tag{B.8}$$

We apply Eq. (B.7) to the square bracket, and rewrite the integral in terms of a shifted variable:

$$\mathscr{T}(\boldsymbol{\delta}\mathbf{x})|j\rangle = \int d^3r'|\mathbf{r}' + \boldsymbol{\delta}\mathbf{x}\rangle\psi_j(\mathbf{r}') = \int d^3r'|\mathbf{r}'\rangle\psi_j(\mathbf{r}' - \boldsymbol{\delta}\mathbf{x})\,. \tag{B.9}$$

We conclude that a translation by $\boldsymbol{\delta}\mathbf{x}$ acts on the wavefunction of a state by substituting its argument $\mathbf{r}'$ with $\mathbf{r}' - \boldsymbol{\delta}\mathbf{x}$.

Translation must conserve the normalization of arbitrary states:

$$\langle j|\mathscr{T}(\boldsymbol{\delta}\mathbf{x})^\dagger\mathscr{T}(\boldsymbol{\delta}\mathbf{x})|j\rangle = \langle j|j\rangle. \tag{B.10}$$

This condition is guaranteed by the requirement that the product

$$\mathscr{T}(\boldsymbol{\delta}\mathbf{x})^\dagger\mathscr{T}(\boldsymbol{\delta}\mathbf{x}) = \mathscr{I}, \tag{B.11}$$

the identity in the Hilbert space. According to Eq. (B.11), $\mathscr{T}(\boldsymbol{\delta}\mathbf{x})$ is a unitary operator, i.e. $\mathscr{T}(\boldsymbol{\delta}\mathbf{x})^\dagger = \mathscr{T}(\boldsymbol{\delta}\mathbf{x})^{-1}$. The following additional properties apply to infinitesimal translations:

$$\mathscr{T}(\boldsymbol{\delta}\mathbf{x})^{-1} = \mathscr{T}(-\boldsymbol{\delta}\mathbf{x})\,, \qquad \mathscr{T}(\mathbf{0}) = \mathscr{I}\,, \qquad \mathscr{T}(\boldsymbol{\delta}\mathbf{x})\mathscr{T}(\boldsymbol{\delta}\mathbf{x}') = \mathscr{T}(\boldsymbol{\delta}\mathbf{x} + \boldsymbol{\delta}\mathbf{x}'). \tag{B.12}$$

The properties of Eqs. (B.11) and (B.12) are satisfied provided that

$$\mathscr{T}(\boldsymbol{\delta}\mathbf{x}) = \mathscr{I} - i\mathbf{K}\cdot\boldsymbol{\delta}\mathbf{x}\,, \tag{B.13}$$

for a suitable "generating" vector $\mathbf{K}$ of self-adjoint operator components $K_{x/y/z}$, and neglecting terms of order $O(|\boldsymbol{\delta}\mathbf{x}|^2)$.

Translations, and their "generators" $\mathbf{K}$ are related to the position operator as follows:

$$\mathbf{R}\,\mathscr{T}(\boldsymbol{\delta}\mathbf{x})|\mathbf{r}'\rangle = \mathbf{R}|\mathbf{r}' + \boldsymbol{\delta}\mathbf{x}\rangle = (\mathbf{r}' + \boldsymbol{\delta}\mathbf{x})|\mathbf{r}' + \boldsymbol{\delta}\mathbf{x}\rangle\,, \tag{B.14}$$

$$\mathscr{T}(\boldsymbol{\delta}\mathbf{x})\mathbf{R}|\mathbf{r}'\rangle = \mathbf{r}'\mathscr{T}(\boldsymbol{\delta}\mathbf{x})|\mathbf{r}'\rangle = \mathbf{r}'|\mathbf{r}' + \boldsymbol{\delta}\mathbf{x}\rangle. \tag{B.15}$$

By subtracting these two relations we have

$$(\mathbf{R}\,\mathscr{T}(\boldsymbol{\delta}\mathbf{x}) - \mathscr{T}(\boldsymbol{\delta}\mathbf{x})\,\mathbf{R})\,|\mathbf{r}'\rangle = \boldsymbol{\delta}\mathbf{x}|\mathbf{r}' + \boldsymbol{\delta}\mathbf{x}\rangle \simeq \boldsymbol{\delta}\mathbf{x}|\mathbf{r}'\rangle, \tag{B.16}$$

where in the last passage we neglect a quantity of order $O(|\boldsymbol{\delta}\mathbf{x}|^2)$. As Eq. (B.16) holds for any $|\mathbf{r}'\rangle$, which form a basis of the Hilbert space, it implies the operatorial relation

$$\mathbf{R}\,\mathscr{T}(\boldsymbol{\delta}\mathbf{x}) - \mathscr{T}(\boldsymbol{\delta}\mathbf{x})\,\mathbf{R} \equiv [\mathbf{R}, \mathscr{T}(\boldsymbol{\delta}\mathbf{x})] = \boldsymbol{\delta}\mathbf{x}\,\mathscr{I}, \tag{B.17}$$

where the $[\cdot,\cdot]$ indicates the commutation relation between two operators. We substitute now Eq. (B.13) into Eq. (B.17), obtaining

$$-\,i\mathbf{R}\,\mathbf{K}\cdot(\boldsymbol{\delta}\mathbf{x}) + i\mathbf{K}\cdot(\boldsymbol{\delta}\mathbf{x})\,\mathbf{R} = \boldsymbol{\delta}\mathbf{x}\,\mathscr{I}\,. \tag{B.18}$$

This relation holds for the 3 components R_u of vector $\mathbf{R}$. We can generate 9 equations by selecting successively the vector $\boldsymbol{\delta}\mathbf{x}$ along the three versors $\hat{\mathbf{x}}_v$:

$$R_u\,K_v - K_v\,R_u \equiv [R_u, K_v] = i\,\delta_{uv}\,\mathscr{I}\,. \tag{B.19}$$

These fundamental commutation relations acquire a central role in QM, once one realizes that $\mathbf{K}$ is related to the *momentum* operator $\mathbf{P}$ by

$$\mathbf{P} = \hbar\mathbf{K}\,, \tag{B.20}$$

where $\hbar = 6.6260 \times 10^{-34}\ \mathrm{m}^2\ \mathrm{kg}\ \mathrm{s}^{-1}$ is the Planck constant. $\mathbf{K}$ is the wave-vector operator. To highlight the role of momentum as the generator of translation, we rewrite Eq. (B.13) as

$$\mathscr{T}(\boldsymbol{\delta}\mathbf{x}) = \mathscr{I} - i\hbar^{-1}\mathbf{P}\cdot\boldsymbol{\delta}\mathbf{x}\,. \tag{B.21}$$

The commutation relations (B.19) are rewritten in terms of momentum as

$$[R_u, P_v] = i\hbar\delta_{uv}\mathscr{I}. \tag{B.22}$$

B.2.1 Compatible and Incompatible Observables

Position and momentum components are examples of non-commuting observables. To understand what this non-commutation implies, we need to clarify first the properties of commuting observables. Suppose that $A\,B$ are operators with $A|a'\rangle = a'|a'\rangle$, $B|b'\rangle = b'|b'\rangle$, and $[A, B] = 0$. Then we prove that if $a' \neq a''$ then $\langle a'|B|a''\rangle = 0$.

This is an immediate consequence of the definition of commutator:

$$0 = \langle a'|\,[A, B]\,|a''\rangle = \langle a'|AB - BA|a''\rangle = (a' - a'')\langle a'|B|a''\rangle\,. \tag{B.23}$$

If all eigenvalues of A are nondegenerate, this means that the basis of eigenstates of A is diagonal for observable B. But even if A has some degenerate eigenvalue, nothing prevents us to diagonalize B in the degenerate Hilbert space, with the result of constructing a basis of *simultaneous eigenstates* of both operators:

$$A|a', b'\rangle = a'|a', b'\rangle\,; \qquad B|a', b'\rangle = b'|a', b'\rangle\,. \tag{B.24}$$

This result can be generalized to any number of commuting observables. The measurement of any of them does not affect the measurement of the others. After the measurement of A the QM system ends up in a common eigenstate $|a', b'\rangle$. Any subsequent sequence of measurements of B and/or A will keep the system in the same eigenstate $|a', b'\rangle$. For this reason, commuting observables are therefore said to be *compatible*.

We have a quite different situation for non commuting observables, such as position and momentum components. First of all, no common basis of eigenkets can exist if $[A, B] \neq 0$. If there was one, then

$$AB|a', b'\rangle = Ab'|a', b'\rangle = a'b'|a', b'\rangle\,; \qquad BA|a', b'\rangle = Ba'|a', b'\rangle = a'b'|a', b'\rangle\,. \tag{B.25}$$

By subtracting these relations, one proves that $[A, B]$ is diagonal and vanishes on a complete basis, which would mean that $[A, B]$ is the null operator, contrary to our assumption. Note that this statement does not prevent A and B from sharing one or several common eigenkets: it just means that these common eigenkets cannot constitute a complete basis.

For any physical state, we define the standard deviation of an arbitrary operator A as

$$\sigma_A^2 = \langle\, (A - \langle A\rangle)^2\rangle = \langle A^2\rangle - \langle A\rangle^2\,, \tag{B.26}$$

where the averages are to be taken over the state under consideration. If the physical state is an eigenstate of A, clearly σ_A^2 vanishes. Otherwise σ_A^2 measures the uncertainty or dispersion to be expected in the outcome of repeated measurements of A realized on that same physical state. Equipped with this definition, it is possible to prove that for any physical state and for any two observables A and B,

$$\sigma_A^2\,\sigma_B^2 \geq \frac{1}{4}\,|\langle[A, B]\rangle|^2\,. \tag{B.27}$$

This *uncertainty relation* expresses the fact that *incompatible observables* cannot both be measured with infinite precision.

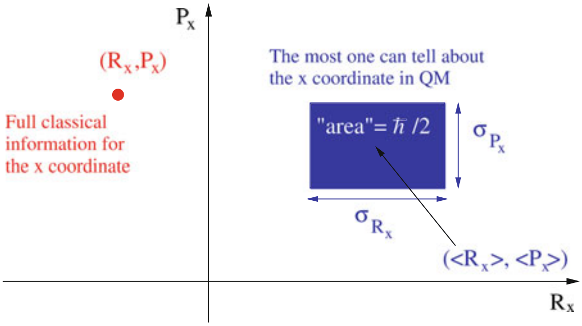


Fig. B.1 A sketch of Heisenberg's uncertainty principle, in the phase-space slice of the $\hat{\mathbf{x}}$ component of the position and momentum of a particle. For classical mechanics (*left*) an individual point in this space represents full information about (R_x, P_x). In contrast, QM (*right*) can only resolve the phase space up to regions of area $\hbar/2$

The example of position and momentum is especially straightforward, since the commutator of corresponding components is a constant $i\hbar$ times the identity operator, see Eq. (B.22). As a consequence of Eq. (B.27), on any physical state *Heisenberg's uncertainty relation*

$$\sigma^2_{R_u}\,\sigma^2_{P_v} \geq \frac{1}{4}\hbar^2\delta_{uv}\,. \tag{B.28}$$

holds. Accordingly, for different components $u \neq v$, e.g. R_x and P_y, which are compatible observables, one can generate states with arbitrarily small uncertainty for both observables. In contrast, for a matching component $u = v$, e.g. R_x and P_x, namely incompatible observables, a state characterized by a small position uncertainty σ_{R_x} has necessarily a large $\sigma_{P_x} \geq \hbar/(2\sigma_{R_x})$, i.e. a largely undefined momentum, and vice versa. This result of QM contrasts with classical mechanics which is based on a phase-space continuous tracking of each particle's both position and momentum, with an in principle arbitrary precision. QM corrects this picture, leading to a fuzzy view of the phase space, with no possibility of identifying sharp points in each pair of conjugate variables such as (R_x, P_x), with a minimum uncertainty "area" of the order of $\hbar/2$, see Fig. B.1.

B.3 The Schrödinger Equation

If one leaves a quantum mechanical system undisturbed, it evolves in time in a predictable manner dictated by the Schrödinger equation associated to its total-energy (or *Hamiltonian*) operator H:

$$i\hbar\frac{d}{dt}|\xi(t)\rangle = H|\xi(t)\rangle\,. \tag{B.29}$$

Here we focus on the common situation of a Hamiltonian H which is constant in time. We have a standard strategy to solve the full time-dependent Schrödinger equation (1.7). This method is based on solving first the "stationary" Schrödinger equation, i.e. the eigenvalue problem

$$H\,|E_j\rangle = E_j|E_j\rangle\,. \tag{B.30}$$

In the resulting basis of *energy eigenkets* $|E_j\rangle$, one then expands the general solution of Eq. (B.29)

$$|\xi(t)\rangle = \sum_j |E_j\rangle\,\langle E_j|\xi(0)\rangle\, e^{-i\,E_j\,t/\hbar}, \tag{B.31}$$

where the time-independent coefficients $\langle E_j|\xi(0)\rangle$ are precisely the complex components of the initial state $|\xi(0)\rangle$ on the adopted basis of energy eigenstates.

There is a remarkable situation, where the system is prepared in a pure energy eigenstate $|\xi(0)\rangle = |E_{\bar{j}}\rangle$. In such a special case, the time evolution involves a single rotating phase factor $\exp(-iE_{\bar{j}}t/\hbar)$. Such an overall phase factor cannot affect the quantum average of any physical quantity: $\langle\xi(t)|Q|\xi(t)\rangle = \langle\xi(0)|Q|\xi(0)\rangle = \langle E_{\bar{j}}|Q|E_{\bar{j}}\rangle$. This property justifies the qualification of the Hamiltonian eigenkets $|E_j\rangle$ as *stationary* states. According to Eq. (B.31), an isolated quantum system prepared in one of its stationary states remains unchanged (up to a trivial rotating phase) forever. As discussed in Sect. 1.2, in real-life time-independent quantum system, excited states would eventually decay under the action of weak interactions with the surrounding environment (e.g. the electromagnetic fields).

An instructive and less trivial condition is realized when the initial state is a linear combination of two energy eigenstates: $|\xi(0)\rangle = a_1\,|E_{j_1}\rangle + a_2\,|E_{j_2}\rangle$. According to Eq. (B.31), the time evolution of this ket is

$$\begin{aligned}|\xi(t)\rangle &= a_1\,e^{-i\,E_{j_1}\,t/\hbar}\,|E_{j_1}\rangle + a_2\,e^{-i\,E_{j_2}\,t/\hbar}\,|E_{j_2}\rangle\\ &= e^{-i\,E_{j_1}\,t/\hbar}\left[a_1\,|E_{j_1}\rangle + a_2\,e^{-i\,(E_{j_2}-E_{j_1})\,t/\hbar}\,|E_{j_2}\rangle\right].\end{aligned}$$

In the final expression, we have extracted an irrelevant overall phase factor, which has no effect on any observable quantity, and highlighted the relative phase of the two contributions in the linear combination. This phase rotates at an angular speed $(E_{j_2} - E_{j_1})/\hbar$ proportional to the energy difference of the two states. At all times that are integer multiple of the rotation period $2\pi\hbar/(E_{j_2} - E_{j_1})$, the evolving ket reproduces its initial condition $|\xi(0)\rangle$ (up to a phase). Note however that when three or more energy eigenkets are involved in the initial condition, no similar return to the initial ket generally occurs.[2]

---

[2] It would occur in the unlikely event that *all* relative frequencies $(E_{j_i} - E_{j_1})/\hbar$ of the eigenkets involved in the initial state are mutually commensurate.

B.4 Symmetry

Symmetries can usually be represented by unitary operators in the Hilbert space: they transform kets into symmetry-modified kets. For example, a mirror reflection plane σ_h takes a ket initially localized at the right side of the plane and moves it to the left side. A second application of this symmetry operator leads back to the original state, thus $\sigma_h \, \sigma_h = \mathscr{I}$. This is not the case with all symmetry operators. For example, a 120° rotation C_3 around a given axis needs being applied 3 times before it leads the kets back to their initial location: $C_3 \, C_3 \, C_3 = \mathscr{I}$. In other cases no number of successive application of a symmetry operation leads back to $\mathscr{I}$, as occurs e.g. with a 1 radian rotation.

The symmetry operators of a given QM system usually form a group, namely a set closed for composition (for any two symmetry operations A and B, also $B\,A$ is a symmetry operation), it contains a neutral element (any system is trivially symmetric under the "do nothing" operator $\mathscr{I}$), and for each A there exists an inverse symmetry operation A^{-1} such that the composition $A^{-1}\,A = A\,A^{-1} = \mathscr{I}$.

Symmetry groups can contain a finite number of elements. For example, the symmetry operations for a water molecule in its equilibrium geometry (see Fig. 3.12b) are 4 (the identity $\mathscr{I}$, a rotation by 180° C_2 around an axis bisecting the $\mathrm{H\hat{O}H}$ angle, and 2 reflection planes: the molecular plane and a perpendicular plane through the C_2 axis).

Other symmetry groups contain an infinite number of elements. A relevant example is the group of the discrete translations of a crystalline solid (the Bravais-lattice translations) described in Sect. 5.1.1. Other important infinite groups are (i) the group of symmetry operations of a linear (e.g. diatomic) molecule, which contains the infinitely many rotations around the molecular axis, (ii) the group of internal symmetry operations of an atom, which includes all possible rotations around its center of mass, and (iii) the group of all continuously many translations of a free particle.

For a given symmetry group, the fact that a QM system has that symmetry is expressed by the fact that any given state $|j\rangle$ has the same energy whether or not it is transformed by any of the group symmetry operators A:

$$\langle j|A^{\dagger} H A|j\rangle = \langle j|H|j\rangle \,. \tag{B.32}$$

As this equation holds for any $|j\rangle$ in the Hilbert space, the corresponding equality of operators must hold:

$$A^{\dagger} H A = H \,, \tag{B.33}$$

or, using the property $A^{\dagger} = A^{-1}$ of unitary operators,

$$HA = AH \,, \text{ or } [H, A] = 0 \,. \tag{B.34}$$

So, all A operators in the group are compatible with H.

In the occasions where A is also an observable (a self-adjoint operator $A^\dagger = A$, which as A is unitary implies $A^{-1} = A^\dagger = A$), as discussed in Sect. B.2.1, H and A can be diagonalized simultaneously. The energy eigenstates are then also symmetry eigenstates. For example, this occurs with the $L \leftrightarrow R$ symmetry operation σ_h of homonuclear dimers: in that system, all one-electron energy eigenstates are either symmetric (bonding) or antisymmetric (antibonding) $L - R$ combinations.

For more complicated symmetry groups, whose elements are not (all) self-adjoint, symmetry provides also a labeling of the energy eigenstates. In these cases, this labeling is given by the irreducible representation of the symmetry group. For the mathematical definition of irreducible representations and how they label the energy eigenstates, the reader is referred to any textbook on group theory, e.g. Ref. [48]. We only need to retain here that many groups have multi-dimensional irreducible representations, thus leading to degenerate energy eigenstates. For example, the degeneracies of the p, d, f, ... electronic one-electron states are related to the 3, 5, 7, ... -dimensional irreducible representations of the full rotational symmetry group of the one-electron Hamiltonian of the atom.

B.5 Variational Methods

Start from an energy eigenstate $|E_j\rangle$ and modify it by adding a small "distortion" $|\delta\rangle$, such that $\langle\delta|\delta\rangle \ll 1$). By taking the normalization of the modified state $|\phi\rangle = |E_j\rangle + |\delta\rangle$ into account, its average energy is:

$$\begin{aligned}
\bar{E}(|E_j\rangle + |\delta\rangle) &= \frac{(\langle E_j| + \langle\delta|)H(|E_j\rangle + |\delta\rangle)}{(\langle E_j| + \langle\delta|)(|E_j\rangle + |\delta\rangle)} \\
&= \frac{\langle E_j|H|E_j\rangle + \langle E_j|H|\delta\rangle + \langle\delta|H|E_j\rangle + \langle\delta|H|\delta\rangle}{\langle E_j|E_j\rangle + \langle E_j|\delta\rangle + \langle\delta|E_j\rangle + \langle\delta|\delta\rangle} \\
&= \frac{E_j + 2\mathrm{Re}\langle\delta|H|E_j\rangle + \langle\delta|H|\delta\rangle}{1 + 2\mathrm{Re}\langle\delta|E_j\rangle + \langle\delta|\delta\rangle} \\
&= \left[E_j + 2E_j\mathrm{Re}\langle\delta|E_j\rangle + \langle\delta|H|\delta\rangle\right] \\
&\quad \times \left[1 - 2\mathrm{Re}\langle\delta|E_j\rangle - \langle\delta|\delta\rangle + 4(\mathrm{Re}\langle\delta|E_j\rangle)^2 + O(|\delta\rangle)^3\right] \\
&= E_j + \langle\delta|H|\delta\rangle - E_j\langle\delta|\delta\rangle + O(|\delta\rangle)^3 \\
&= E_j + O(|\delta\rangle)^2 .
\end{aligned}$$

This result illustrates that changes to the mean energy of the modified state are at least quadratic in the norm of the perturbing ket $|\delta\rangle$. This argument holds for the ground state, which is a minimum of the energy as a function of the ket; excited states are stationary saddle points.

Given a set of states $\{|\phi_j\rangle\}$ one can build linear combinations $|\phi\rangle = \sum_l c_l |\phi_l\rangle$ of these states. Assume that these states $\{|\phi_j\rangle\}$ are elements of a Hilbert space where an energy operator H is defined. The best approximation to the energy eigenstates $|E_j\rangle$ and eigenenergies E_j is realized by tuning the coefficients $\{c_l\}$ in such a way that they implement the stationary property discussed above: the average energy $\bar{E}(|\phi\rangle) = \bar{E}(\{c_l\}) = \langle\phi|H|\phi\rangle/\langle\phi|\phi\rangle$ must have null linear-order variation. $\bar{E}$ should vary at least quadratically as the coefficients $\{c_l\}$ are displaced away from their optimal values. To determine the values of the coefficients making $\bar{E}$ *stationary*, we impose that its gradient against variations of c_j^* vanishes:

$$0 = \frac{\partial \bar{E}(|\phi\rangle)}{\partial c_j^*} = \frac{\langle\phi_j|H|\phi\rangle}{\langle\phi|\phi\rangle} - \frac{\langle\phi|H|\phi\rangle\,\langle\phi_j|\phi\rangle}{(\langle\phi|\phi\rangle)^2} = \sum_l \frac{\langle\phi_j|H|\phi_l\rangle}{\langle\phi|\phi\rangle}\,c_l - \bar{E}\sum_l \frac{\langle\phi_j|\phi_l\rangle}{\langle\phi|\phi\rangle}\,c_l\,. \tag{B.35}$$

Introduce the shorthand $H_{jl} = \langle\phi_j|H|\phi_l\rangle$ for the matrix element, and $B_{jl} = \langle\phi_j|\phi_l\rangle$ for the overlap ($B_{jl} = \delta_{jl}$ if orthonormal states $\{|\phi_j\rangle\}$ are selected). In this notation Eq. (B.35) becomes:

$$\sum_l H_{jl}\,c_l = \bar{E}\sum_l B_{jl}\,c_l\,. \tag{B.36}$$

This variational method transforms the original abstract Schrödinger problem (B.30) into an algebraic problem, namely the generalized secular problem of the calculation of the approximate eigenvalues $\bar{E}$ and eigenvectors of a matrix $\{H_{jl}\}$ relative to the overlap matrix $\{B_{jl}\}$. Note that, given a Hilbert subspace generated by N_ϕ states $|\phi_j\rangle$, at most N_ϕ eigenenergies and eigenstates can be estimated with this method. In particular, the lowest eigenvalue of the matrix problem (B.36) is an upper bound of the exact ground-state energy: $\bar{E}_0 \geq E_0$.

The form (B.36) is standard and ready to translate into a computer implementation. In several applications, an orthonormal basis ($B_{jl} = \delta_{jl}$) is adopted for conceptual simplicity and numerical convenience.

B.5.1 One State

An especially simple application of the variational method is obtained for $N_\phi = 1$. In this case, the variational inequality $H_{00} \equiv \bar{E}_0 \geq E_0$ states the trivial fact that the average energy of an arbitrary state is greater than or equal to the ground-state energy. The single trial state $|\phi\rangle$ is often made depend on some tunable parameters $a_1, a_2, \ldots$ When parameters are available, they can be “optimized” with the target of minimizing the average energy $\bar{E}_0 = H_{00}$:

$$E_0 \leq \min_{a_1,a_2,\ldots} \bar{E}_0 \equiv \min_{a_1,a_2,\ldots} H_{00} \equiv \min_{a_1,a_2,\ldots} \langle\phi(a_1,a_2,\ldots)|\,H\,|\phi(a_1,a_2,\ldots)\rangle\,, \tag{B.37}$$

thus realizing the best approximation (in this class of parameterized states) to the true ground state.

The HF method of Sect. 2.2.4 is an example of this single-state approach. The parameters being optimized in that case are the single-electron wave functions themselves, which parameterize the many-body state.

B.5.2 Two States

The solution of Eq. (B.36) for two states $|\phi_1\rangle = |L\rangle$ and $|\phi_2\rangle = |R\rangle$ is especially instructive. For simplicity assume that these states are orthonormal, so that the overlap matrix $B_{jl} = \langle\phi_j|\phi_l\rangle = \delta_{jl}$ is the identity. The eigenenergies and eigenstates of the full problem are approximated by the eigenvalues and eigenvectors of the 2 × 2 matrix

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \equiv \begin{pmatrix} \langle L|H|L\rangle & \langle L|H|R\rangle \\ \langle R|H|L\rangle & \langle R|H|R\rangle \end{pmatrix} = \begin{pmatrix} E_L & -\Delta \\ -\Delta^* & E_R \end{pmatrix}. \tag{B.38}$$

The eigenvalues of the matrix (B.38) are

$$\bar{E}_{\frac{1}{2}} = \frac{E_L + E_R}{2} \mp \sqrt{\left(\frac{E_L - E_R}{2}\right)^2 + |\Delta|^2}\,. \tag{B.39}$$

The corresponding ground $|\bar{E}_1\rangle$ and excited $|\bar{E}_2\rangle$ eigenkets can be written as

$$\left|\bar{E}_{\frac{1}{2}}\right\rangle = \frac{\Delta}{|\Delta|}\sqrt{\frac{1}{2}\left(1 \mp \frac{u}{\sqrt{1+u^2}}\right)}\,|L\rangle \pm \sqrt{\frac{1}{2}\left(1 \pm \frac{u}{\sqrt{1+u^2}}\right)}\,|R\rangle\,, \tag{B.40}$$

assuming $E_L \geq E_R$, and with

$$u = \frac{E_L - E_R}{2|\Delta|}\,. \tag{B.41}$$

measuring the relative importance of the diagonal energy difference to the off-diagonal coupling strength.

The eigenvalues are sketched in Fig. B.2. We see that the two eigenenergies are centered around $(E_L + E_R)/2$, due to the conservation of the matrix trace upon diagonalization. The eigenenergies never get closer than the smaller of $(E_R - E_L)$ and $|2\Delta|$. The off-diagonal coupling Δ induces a sort of mutual "repulsion" between the energy levels.

In the special limit $E_L = E_R$, i.e. $u = 0$, the two eigenvalues acquire a minimum separation $|2\Delta|$; the corresponding kets turn into symmetric $|\bar{E}_1\rangle = 2^{-1/2}(|L\rangle + |R\rangle)$ and antisymmetric $|\bar{E}_2\rangle = 2^{-1/2}(|L\rangle - |R\rangle)$ combinations of the original states. For increasing u, the eigenkets (B.40) deviate more and more from these simple

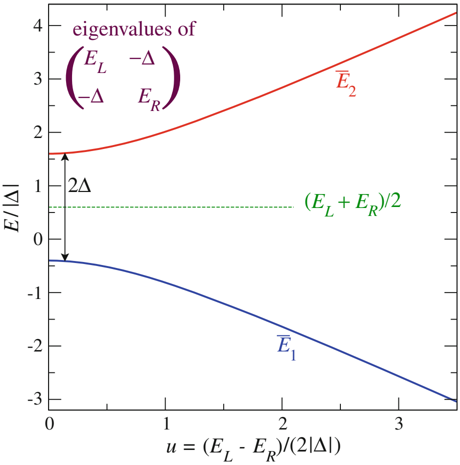


Fig. B.2 The eigenvalues, Eq. (B.39), of the 2 × 2 matrix (B.38), as a function of the asymmetry ratio u. The symmetric case $u = 0$ describes the splitting of degenerate levels by a off-diagonal "perturbation". This splitting determines, e.g., the bonding-antibonding splitting in the electronic structure of homonuclear diatomic molecules. For large $u \gg 1$, $E_b \simeq E_R$ and $E_a \simeq E_L$. This limit describes, e.g., the electronic bonding/antibonding states of strongly ionic dimers

symmetric and antisymmetric combinations: $|b\rangle$ acquires a prevalent $|R\rangle$ character, while $|a\rangle$ acquires a mainly $|L\rangle$ character.

B.6 The Schrödinger Equation in Real Space

The variational method outlined above provides also a useful reformulation of the Schrödinger equation, by adopting the basis of position eigenstates $|\mathbf{r}'\rangle$ for the set of kets $|\phi_l\rangle$. Because for the translational degrees of freedom the set of all $|\mathbf{r}'\rangle$ is complete, this method yields an exact mapping of the Schrödinger equation (B.30) for the translational degrees of freedom. This mapping to a differential equation for the wavefunction $\psi(\mathbf{r}')$ takes the name of "position representation".

To obtain this mapping, we need to represent the momentum operator on the position basis. Applying Eq. (B.21) on an arbitrary state $|\alpha\rangle$, we have

$$\begin{aligned}\left(\mathscr{I} - i\hbar^{-1}\mathbf{P}\cdot\boldsymbol{\delta}\mathbf{x}\right)|\alpha\rangle &= \mathscr{T}(\boldsymbol{\delta}\mathbf{x})|\alpha\rangle = \mathscr{T}(\boldsymbol{\delta}\mathbf{x})\int d^3r'\,|\mathbf{r}'\rangle\langle\mathbf{r}'|\alpha\rangle \\ &= \int d^3r'\,\mathscr{T}(\boldsymbol{\delta}\mathbf{x})|\mathbf{r}'\rangle\langle\mathbf{r}'|\alpha\rangle = \int d^3r'\,|\mathbf{r}' + \hat{\mathbf{x}}\,\delta x\rangle\langle\mathbf{r}'|\alpha\rangle \\ &= \int d^3r'\,|\mathbf{r}'\rangle\langle\mathbf{r}' - \hat{\mathbf{x}}\,\delta x|\alpha\rangle\end{aligned}$$

$$= \int d^3r' \, |\mathbf{r}'\rangle \left(\langle \mathbf{r}'|\alpha\rangle - \boldsymbol{\delta}\mathbf{x} \cdot \nabla_{r'} \langle \mathbf{r}'|\alpha\rangle\right)$$

$$= |\alpha\rangle - \boldsymbol{\delta}\mathbf{x} \cdot \int d^3r' \, |\mathbf{r}'\rangle \nabla_{r'} \langle \mathbf{r}'|\alpha\rangle.$$

By identifying the terms proportional to $\boldsymbol{\delta}\mathbf{x}$, we conclude that

$$\mathbf{P}|\alpha\rangle = -i\hbar \int d^3r' \, |\mathbf{r}'\rangle \nabla_{r'} \langle \mathbf{r}'|\alpha\rangle = -i\hbar \int d^3r' \, |\mathbf{r}'\rangle \nabla_{r'} \psi_\alpha(\mathbf{r}'). \tag{B.42}$$

This expression indicates that the wavefunction of a state after the momentum operator has acted on it is $-i\hbar \times$ the gradient of the wavefunction $\psi_\alpha(\mathbf{r}')$ of the original state. In this sense, the position representation of the momentum operator is $-i\hbar \, \nabla_{r'}$.

The following expressions for the matrix elements of $\mathbf{P}$ are direct consequences of Eq. (B.42):

$$\langle \mathbf{r}'|\mathbf{P}|\mathbf{r}''\rangle = -i\hbar \nabla_{r'} \delta(\mathbf{r}' - \mathbf{r}'') \tag{B.43}$$

$$\langle \beta|\mathbf{P}|\alpha\rangle = \int d^3r' \, \psi_\beta^*(\mathbf{r}') \, (-i\hbar \nabla_{r'}) \, \psi_\alpha(\mathbf{r}') \, . \tag{B.44}$$

For a single particle of mass m under the action of conservative forces described by a potential energy function $V(\mathbf{r})$, the Hamiltonian is

$$H = H_{\text{kin}} + V(\mathbf{R}) = \frac{|\mathbf{P}|^2}{2m} + V(\mathbf{R}) \, . \tag{B.45}$$

The expression (B.43) for the momentum matrix operator allows us to recognize that on the position basis the matrix of the $|\mathbf{P}|^2$ operator is diagonal:

$$\langle \mathbf{r}'| \, |\mathbf{P}|^2 |\mathbf{r}''\rangle = -\hbar^2 \nabla_{r'}^2 \, \delta(\mathbf{r}' - \mathbf{r}''). \tag{B.46}$$

Also the potential-energy term is of course diagonal on this basis:

$$\langle \mathbf{r}'|V(\mathbf{R})|\mathbf{r}''\rangle = V(\mathbf{r}') \, \delta(\mathbf{r}' - \mathbf{r}'') \, . \tag{B.47}$$

With these results in mind, evaluate now the matrix elements in Eq. (B.36) on the position basis: $|\phi_j\rangle = |\mathbf{r}'\rangle$ and $|\phi_l\rangle = |\mathbf{r}''\rangle$. Due to the Dirac deltas in Eqs. (B.46) and (B.47), the $\sum_l$ (here an $\int d^3r''$) drops out, resulting in a substantial simplification. Therefore, the Schrödinger equation (B.30) maps to

$$\left[-\frac{\hbar^2}{2m} \, \nabla_{r'}^2 + V(\mathbf{r}') \right] \psi_{\mathscr{E}}(\mathbf{r}') = \mathscr{E} \, \psi_{\mathscr{E}}(\mathbf{r}') \, . \tag{B.48}$$

This is a second-order differential secular equation in 3D space. Simpler 1D and 2D versions can be formulated whenever the motion in the remaining dimensions is either impossible or trivial. For several interacting particles moving simultaneously, under the action of a many-body Hamiltonian

$$H = H_{\text{kin}} + V(\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3, \dots) = \sum_\alpha \frac{|\mathbf{P}_\alpha|^2}{2m} + V(\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3, \dots)\,, \tag{B.49}$$

the equation in real space generalizes to

$$\left[-\frac{\hbar^2}{2m}\sum_\alpha \nabla^2_{r'_\alpha} + V(\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}'_3, \dots)\right]\psi_E(\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}'_3, \dots) = E\,\psi_E(\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}'_3, \dots)\,. \tag{B.50}$$

For most practical problems, the real-space formulation (B.48) and (B.50) is far too complicated to solve analytically. In these cases one usually goes back to approximate numerical methods based on the mapping of Eq. (B.30) onto an algebraic matrix problem, as described in Sect. B.5. However, the real-space formulation of the Schödinger problem can be solved exactly in a few simple cases.

B.7 Simple Solutions of the Schrödinger Equation

In the following we report three simple but physically fundamental examples. A few other examples, e.g. the one-electron atom, are mentioned elsewhere in the present book.

B.7.1 A Free Particle

When in Eq. (B.48) the potential energy $V(\mathbf{r}') = 0$, the motion of the particle is free. One can separate the motion in the three Cartesian directions by decomposing the solution

$$\psi_{\mathscr{E}}(\mathbf{r}') = \psi_{\mathscr{E}_x}(r'_x)\,\psi_{\mathscr{E}_y}(r'_y)\,\psi_{\mathscr{E}_z}(r'_z)\,, \tag{B.51}$$

with $\mathscr{E}_u$ representing the contribution of the u-component of the motion to the total kinetic energy $\mathscr{E} = \mathscr{E}_x + \mathscr{E}_y + \mathscr{E}_z$. The equation for each component is

$$-\frac{\hbar^2}{2m}\nabla^2_{r'_u}\,\psi_{\mathscr{E}_u}(r'_u) = \mathscr{E}_u\,\psi_{\mathscr{E}_u}(r'_u)\,. \tag{B.52}$$

Any kind of harmonic-type functions such as, e.g., $\psi(r'_u) = \exp(\kappa r'_u)$ or $\psi(r'_u) = \sin(\kappa r'_u)$ solve Eq. (B.52) for a suitable $\mathscr{E}_u$. However, not all solutions are equally acceptable: in these two examples, whenever $\Re(\kappa) \neq 0$ the exponential one would diverge at large r'_u, and for $\Im(\kappa) \neq 0$ the trigonometric one would have similar defects. In practice, solutions of the type

$$\psi_{k_u}(r'_u) = (2\pi)^{-1/2} \exp(i\, k_u\, r'_u) \tag{B.53}$$

are usually adopted, for real k_u. The factor $(2\pi)^{-1/2}$ yields a convenient normalization for a wavefunction defined in the whole range $-\infty < r'_u < \infty$. In such a case, all real values of k_u are possible, and provide an energy contribution $\mathscr{E}_u = \hbar^2 k_u^2/(2m)$, as one can immediately verify by inserting the solution (B.53) into Eq. (B.52).

The vector $\mathbf{k}$ of components $\{k_x, k_y, k_z\}$ represents the eigenvalue of the vector operator $\mathbf{K}$ introduced in Eq. (B.13). Indeed, the states defined by the plane-wave function of Eq. (B.53) are eigenstates $|\mathbf{k}\rangle$ of $\mathbf{K}$ and therefore also of the translation operator $\mathscr{T}(\boldsymbol{\delta}\mathbf{x})$. According to Eq. (B.20), these $|\mathbf{k}\rangle$ states are also eigenstates of the momentum operator $\mathbf{P}$ with eigenvalue $\mathbf{p}' = \hbar\mathbf{k}$. The ordinary relations of energy and momentum hold for these eigenvalues:

$$\mathscr{E} = \mathscr{E}_x + \mathscr{E}_y + \mathscr{E}_z = \frac{\hbar^2 k_x^2}{2m} + \frac{\hbar^2 k_y^2}{2m} + \frac{\hbar^2 k_z^2}{2m} = \frac{p'^2_x}{2m} + \frac{p'^2_y}{2m} + \frac{p'^2_z}{2m} = \frac{|\mathbf{p}'|^2}{2m}. \tag{B.54}$$

It is often convenient to express the free-particle wavefunctions in a finite—rather than infinite—region of space, of size $V = L \times L \times L$. Periodic boundary conditions are usually applied across this cube. As a consequence, not all values of wave vector are allowed, but only those with components

$$k_u = \frac{2\pi}{L}\, n_u \qquad n_u = 0, \pm 1, \pm 2, \pm 3, \dots \tag{B.55}$$

The associated plane-wave eigenfunctions are normalized as follows:

$$\psi_{\mathbf{k}}(\mathbf{r}) = L^{-3/2} \exp(i\, \mathbf{k} \cdot \mathbf{r}). \tag{B.56}$$

The corresponding translational kinetic energy takes on the discrete values

$$\mathscr{E}_{\mathbf{n}} = \frac{|\mathbf{p}'_{\mathbf{n}}|^2}{2m} = \frac{\hbar^2 |\mathbf{k}_{\mathbf{n}}|^2}{2m} = \frac{(2\pi\hbar)^2}{2mL^2}\,(n_x^2 + n_y^2 + n_z^2). \tag{B.57}$$

In the limit of infinite size $L \to \infty$ the discrete kinetic-energy eigenvalues turn into the same continuum of positive energies as as those of Eq. (B.54).

B.7.2 A Particle in an Infinitely Deep Square Well

A particle confined in a finite region of space is particularly illuminating the consequence of the Heisenberg uncertainty relation on the energy spectrum.

Consider for simplicity a particle confined in 1D to an interval by the potential energy

$$V(r_x) = \begin{cases} 0 & \text{if } 0 \le r_x \le L \\ +\infty & \text{elsewhere} \end{cases} . \tag{B.58}$$

Inside the confinement region, the particle moves freely: its wavefunction follows Eq. (B.52). The confining potential forces the wave function to vanish at $r_x \le 0$ and $r_x \ge L$, or else the infinite repulsion would make $\mathscr{E}_x$ diverge. This requirement imposes a vanishing boundary condition to the eigenfunctions of Eq. (B.52). The general solution

$$\psi_{k_x}(r_x) = \sqrt{\frac{2}{L}}\, \sin(k_x\, r_x)\,, \quad k_x = \frac{\pi\, n}{L}\,, \quad n = 1, 2, 3, \ldots \tag{B.59}$$

exhibits $n - 1$ nodes, where it changes sign.

The corresponding (kinetic) energy eigenvalues are

$$\mathscr{E}_n = \frac{\hbar^2 k_x^2}{2m} = \frac{\pi^2 \hbar^2\, n^2}{2mL^2}\,, \quad n = 1, 2, 3, \ldots . \tag{B.60}$$

Compare this result with Eq. (4.39): note in particular that the ground-state energy of the free particle vanishes, while that of the confined particle, obtained by substituting $n = 1$ in Eq. (B.60), is nonzero. Its value increases with the inverse of the particle mass and the inverse square of the confinement size L. This a fundamental consequence of Heisenberg's uncertainty relation: the more restrictively a quantum particle is confined in position space, the fuzzier its momentum becomes, with a resulting larger and larger average kinetic energy.

B.7.3 The Linear Harmonic Oscillator

Another basic problem of mechanics is that of a mass attached to a fixed point through an elastic spring, which we address for simplicity in 1D. The restoring force $-Kr_x$ is represented by the potential energy

$$V(r_x) = \frac{1}{2} K r_x^2. \tag{B.61}$$

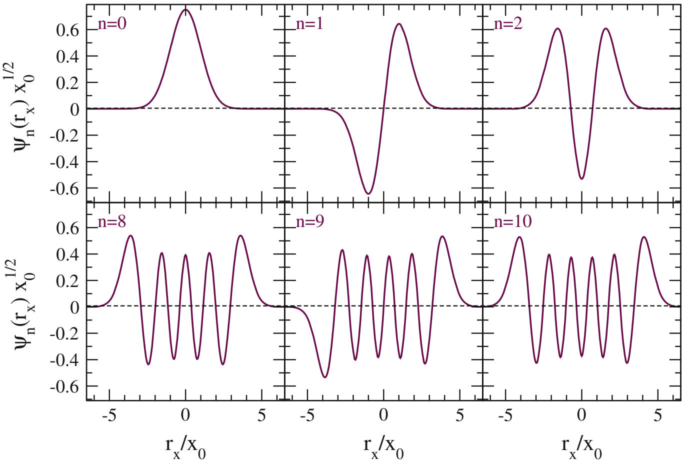


Fig. B.3 The ground-state and a few excited eigenfunctions of the linear harmonic oscillator

As it is straightforward to verify by direct substitution, the eigenfunctions of the Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\,\nabla^2_{r_x}+\frac{1}{2}Kr_x^2\right)\psi_{\mathscr{E}}(r_x)=\mathscr{E}\;\psi_{\mathscr{E}}(r_x) \tag{B.62}$$

are

$$\psi_n(r_x)=\left(\frac{1}{\pi^{1/2}2^n n!\,x_0}\right)^{1/2}\,H_n\left(\frac{r_x}{x_0}\right)\,\exp\left(-\frac{r_x^2}{2x_0^2}\right),\quad n=0,1,2,3,\ldots, \tag{B.63}$$

where

$$x_0=\hbar^{1/2}\,(mK)^{-1/4}\,,\qquad H_n(\xi)=(-1)^n\exp\left(\xi^2\right)\frac{d^n}{d\xi^n}\exp\left(-\xi^2\right). \tag{B.64}$$

The functions $H_n(\xi)$ are (Hermite) polynomials of degree n, e.g. $H_0(\xi)=1, H_1(\xi)=2\xi, H_2(\xi)=4\xi^2-2$, … As illustrated in Fig. B.3, $\psi_n(r_x)$ exhibits n nodes, where $H_n(r_x/x_0)$ changes sign.

The corresponding energy eigenvalues

$$\mathscr{E}_n=\hbar\omega_0\left(n+\frac{1}{2}\right)=\hbar\sqrt{\frac{K}{m}}\left(n+\frac{1}{2}\right) \tag{B.65}$$

form an equally spaced ladder of spacing $\hbar\omega_0$. The energy $\mathscr{E}_n$ of each eigenstate is composed of 50 % kinetic plus 50 % potential contributions.

The ground-state position probability distribution $P_0(r_x) = |\psi_0(r_x)|^2$ is a Gaussian of standard deviation $2^{-1/2}x_0$, measuring the amplitude of the zero-point motion. Like for the infinite square well, the "zero-point" ground-state energy $\mathscr{E}_0 = \hbar\omega_0/2$ is positive, as a consequence of confinement and Heisenberg's uncertainty relation. The 50 % kinetic and 50 % potential energy composition of $\mathscr{E}_0$ contrasts the purely-kinetic ground-state energy $\mathscr{E}_1$ of the infinite square well, Eq. (B.60).

The harmonic oscillator is a fundamental model for several physical phenomena, including

- the vibrational degree of freedom of a diatomic molecule, where v is used for the quantum number n, see Sect. 3.3;
- each normal mode of vibration of a polyatomic molecule, see Fig. 3.12;
- each normal mode of vibration of a solid, see Sect. 5.3;
- each normal oscillation mode of the electromagnetic fields in a cavity, see Sect. 4.3.2.2.

B.8 Angular Momentum

Similar to linear momentum **P**, which generates translations, see Sect. B.2, angular momentum generates rotations. The orbital angular momentum operator $\mathbf{L} = \mathbf{R} \times \mathbf{P}$ generates the rotations of the position degrees of freedom of a particle around the origin of the reference system. Starting from the commutation relations for the **R** and **P** operators, it is straightforward to realize that $[L_x, L_y] = i\hbar L_z$, and analogous relations obtained by cyclic permutations of the components. These commutation relations are fundamental in nature, and can be deduced directly for the total angular momentum **J** from requiring that it is related to the operator $\mathscr{D}(\hat{\mathbf{n}}, \delta\phi)$ implementing a rotation by an infinitesimal angle $\delta\phi$ around the direction $\hat{n}$, by

$$\mathscr{D}(\hat{\mathbf{n}}, \delta\phi) = \mathscr{I} - i\hbar^{-1}\mathbf{J} \cdot \hat{\mathbf{n}}\,\delta\phi\,. \tag{B.66}$$

One realizes that $\mathbf{J} \equiv \mathbf{L}$ for a spinless particle. For particles with spin instead, a rotation must include spin as well, with the result that an extra contribution has to be included in the *total angular momentum*

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{B.67}$$

of a particle carrying spin, such as an electron. Spin must of course fulfill the same fundamental commutation relations $[S_x, S_y] = i\hbar S_z$, etc. as the L_u components. Moreover, spin acts in a space different from the one of translations, therefore $[L_u, S_v] = 0$. For systems formed by several particles the rotations of all must be included, so that

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 + \mathbf{J}_3 + \cdots \tag{B.68}$$

The angular momentum commutation relations

$$[J_x,\ J_y] = i\,\hbar\,J_z\,,\ \text{etc.} \tag{B.69}$$

are all that it takes to determine the properties (eigenvalues, eigenvectors, matrix elements) of the angular momentum operators, irrespective of their spin or orbital nature [17]. One can show that the operator $|\mathbf{J}|^2$ commutes with any component operator, e.g. J_z. As a consequence, $|\mathbf{J}|^2$ and J_z are compatible observables and can be diagonalized simultaneously, see Sect. B.2.1. Call $|j, m_j\rangle$ the basis of common eigenstates of $|\mathbf{J}|^2$ and J_z. The expressions for their eigenvalues are [17]:

$$|\mathbf{J}|^2|j, m_j\rangle = \hbar^2 j(j+1)\,|j, m_j\rangle\,, \quad \text{with}\ j = 0,\ {}^1\!/\!{}_2,\ 1,\ {}^3\!/\!{}_2,\ 2,\ {}^5\!/\!{}_2, \ldots \tag{B.70}$$

$$J_z|j, m_j\rangle = \hbar m_j\,|j, m_j\rangle\,, \quad \text{with}\ m_j = -j,\ -j+1,\ \ldots\ j-1,\ j\,. \tag{B.71}$$

For given j, the projection quantum number m_j can take one of the $2j+1$ values listed above.

B.8.1 The Coupling of Angular Momenta

A common problem arises from the necessity of combining several angular momenta, e.g. those carried by several particles or by the orbital and spin degrees of freedom of the same particle. The main question is: what are the allowed eigenvalues of the total angular momentum j, for given values of the coupled angular momenta?

Here we provide the answer to this problem for the addition of two angular momenta only. Say you want to add an orbital angular momentum $\mathbf{L}$ and a spin one $\mathbf{S}$. The starting observation is that one can identify two sets of 4 mutually commuting operators ($|\mathbf{L}|^2$, $|\mathbf{S}|^2$, L_z, S_z) and ($|\mathbf{L}|^2$, $|\mathbf{S}|^2$, $|\mathbf{J}|^2$, J_z). Each of these sets can be diagonalized simultaneously to produce a basis of orthonormal states. Basis one and basis two, obtained diagonalizing the first and second set respectively, do not coincide. However, observe that two out of four operators coincide $|\mathbf{L}|^2$ and $|\mathbf{S}|^2$. They can then be taken diagonal in both bases. It is then possible to fix the corresponding quantum numbers l and s.

In basis one, of states $|l, s, m_l, m_s\rangle$, L_z and S_z are diagonal. For fixed l and s, this basis is formed by $d = (2l+1)\cdot(2s+1)$ states labeled by all possible combinations of the allowed values $m_l = -l,\ -l+1,\ \ldots\ l-1,\ l$ and $m_s = -s,\ -s+1,\ \ldots\ s-1,\ s$ for the individual $\hat{\mathbf{z}}$-projections. Alternatively, one can select the second basis of states $|l, s, j, m_j\rangle$, with diagonal $|\mathbf{J}|^2$ and J_z, to span this d-dimensional space of states. The coupling of these angular momenta provides the change of basis from basis one to basis two.

To identify the $|l, s, j, m_j\rangle$ states we need to identify the values of j compatible with the given l and s. The commutation relations are sufficient to answer completely this question [17] and also to express the “coupled states” $|l, s, j, m_j\rangle$ in terms of the

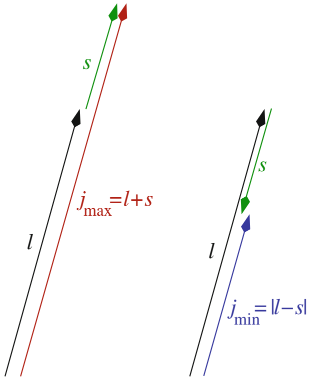


Fig. B.4 An intuitive mnemonic sketch for the rule of angular-momentum composition, Eq. (B.72)

original L_z and S_z eigenstates $|l, s, m_l, m_s\rangle$. For our purposes, it suffices to retain the main result of this instructive exercise, namely that the allowed values for j are

$$j = |l - s|,\ |l - s| + 1,\ \ldots\ l + s - 1,\ l + s\,. \tag{B.72}$$

The extremal values recall the classical picture of vector composition (Fig. B.4); the discrete values are characteristic of the QM of angular momentum. One should check that the number d of states in basis one and two is the same, i.e. that

$$(2l + 1) \cdot (2s + 1) = 2(|l - s|) + 1\ +\ 2(|l - s| + 1) + 1\ +\ \cdots \\ + 2(l + s - 1) + 1\ +\ 2(l + s) + 1\,.$$

The coupling of angular-momentum eigenstates is realized by a unitary transformation in this d-dimensional space, i.e. by multiplication of the basis states by a unitary $d \times d$ matrix:

$$|l, s, j, m_j\rangle = \sum_{m_l, m_s} C^{j\, m_j}_{l\, m_l\ s\, m_s} |l, s, m_l, m_s\rangle\,, \tag{B.73}$$

where the $C^{j\, m_j}_{l\, m_l\ s\, m_s}$ numbers (named Clebsch-Gordan coefficients, taken real by convention, and tabulated in books [49, 50]) provide the coefficients for the linear combinations forming the transformed basis.

Unless some physical coupling mechanism is present, either basis is equally suitable to describe a particle's angular+spin dynamics. The difference is that the $|m_l,\ m_s\rangle$ basis (here we drop the fixed l and s labels) emphasizes the invariance against separate position-space and spin-space rotation, while the $|j,\ m_j\rangle$ basis emphasizes the invariance for global rotations (equal rotations for position and spin).

Consider the concrete example of an electron with its spin. With the basic nonrelativistic Hamiltonian (1.1) all the d states within the (l,s) multiplet have the same energy, and this degeneracy occurs whether we describe them in terms of the $|m_l,\ m_s\rangle$

basis or of the $|j,\ m_j\rangle$ basis. Taking relativistic interactions into account, which are invariant over global rotations, not separate orbital and spin rotations, the coupled $|j,\ m_j\rangle$ basis becomes the basis of choice. Let us see then how the orbital angular momentum l of one electron combines with its spin $s = 1/2$. Rule (B.72) attributes to j either one value $j = s = 1/2$ (when $l{=}0$, i.e. s states), or two values $j = l \pm 1/2$ (for all other $l = 1,\ 2,\ 3\ldots$, i.e. p, d, f… states). For s states, the transformation matrix from the $|m_l = 0,\ m_s\rangle$ basis to the $|j{=}1/2,\ m_j\rangle$ basis is trivially the 2×2 identity $C^{1/2\, m_j}_{0\,0\ 1/2\, m_s} = \delta_{m_j\, m_s}$, since, for $l = 0$, $\mathbf{J}$ coincides with $\mathbf{S}$. Likewise, for $l \geq 1$, the spin-orbital "maximally aligned" components

$$\left|j{=}l{+}\frac{1}{2},\ m_j{=}\pm\left(l{+}\frac{1}{2}\right)\right\rangle = \left|m_l{=}\pm l,\ m_s{=}\pm\frac{1}{2}\right\rangle.$$

Each of the remaining coupled states is expressed in terms of two uncoupled states only, namely those whose J_z component match:

$$\left|j{=}l+\frac{1}{2},\ m_j\right\rangle = c_1\ \left|m_l{=}m_j{+}\frac{1}{2},\ m_s{=}{-}\frac{1}{2}\right\rangle + c_2\ \left|m_l{=}m_j{-}\frac{1}{2},\ m_s{=}{+}\frac{1}{2}\right\rangle.$$

The orthogonal ket is

$$\left|j{=}l{-}\frac{1}{2},\ m_j\right\rangle = c_2\ \left|m_l{=}m_j{+}\frac{1}{2},\ m_s{=}{-}\frac{1}{2}\right\rangle - c_1\ \left|m_l{=}m_j{-}\frac{1}{2},\ m_s{=}{+}\frac{1}{2}\right\rangle,$$

where

$$c_1 = C^{l+\frac{1}{2}\, m_j}_{l\, m_j+\frac{1}{2}\ \frac{1}{2}\ -\frac{1}{2}} = \sqrt{\frac{l+\frac{1}{2}-m_j}{2l+1}} \quad \text{and} \quad c_2 = C^{l+\frac{1}{2}\, m_j}_{l\, m_j-\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}} = \sqrt{\frac{l+\frac{1}{2}+m_j}{2l+1}}$$

are the matrix elements of the basis-change transformation. Clearly, $c_1^2 + c_2^2 = 1$ must hold for this transformation to be unitary.

Even more concretely, for a p ($l = 1$) orbital triplet, the explicit transformation matrix between the $|m_l,\ m_s\rangle$ basis and the coupled $|j,\ m_j\rangle$ basis involves a 6×6 matrix as follows:

$$\begin{pmatrix} |j=\frac{1}{2},\ m_j=\frac{1}{2}\rangle \\ |j=\frac{1}{2},\ m_j=-\frac{1}{2}\rangle \\ |j=\frac{3}{2},\ m_j=\frac{3}{2}\rangle \\ |j=\frac{3}{2},\ m_j=\frac{1}{2}\rangle \\ |j=\frac{3}{2},\ m_j=-\frac{1}{2}\rangle \\ |j=\frac{3}{2},\ m_j=-\frac{3}{2}\rangle \end{pmatrix} = \begin{pmatrix} 0 & -\sqrt{\frac{1}{3}} & 0 & \sqrt{\frac{2}{3}} & 0 & 0 \\ 0 & 0 & -\sqrt{\frac{2}{3}} & 0 & \sqrt{\frac{1}{3}} & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{\frac{2}{3}} & 0 & \sqrt{\frac{1}{3}} & 0 & 0 \\ 0 & 0 & \sqrt{\frac{1}{3}} & 0 & \sqrt{\frac{2}{3}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} |m_l=1,\ m_s=+\frac{1}{2}\rangle \\ |m_l=0,\ m_s=+\frac{1}{2}\rangle \\ |m_l=-1,\ m_s=+\frac{1}{2}\rangle \\ |m_l=1,\ m_s=-\frac{1}{2}\rangle \\ |m_l=0,\ m_s=-\frac{1}{2}\rangle \\ |m_l=-1,\ m_s=-\frac{1}{2}\rangle \end{pmatrix},$$

where we list the $j = 1 - 1/2 = 1/2$ doublet followed by the $j = 1 + 1/2 = 3/2$ quartet. In a pictorial "box" notation where $\begin{array}{|c|c|c|}\hline & & \uparrow \\ \hline & & \\ \hline\end{array}$ stands for the uncoupled basis state $|m_l = 1,\ m_s = +1/2\rangle$, e.g., the two bottom rows of the matrix relation above can be written as $\left|j = 3/2,\ m_j = -1/2\right\rangle = \sqrt{\frac{1}{3}}\begin{array}{|c|c|c|}\hline \uparrow & & \\ \hline & & \\ \hline\end{array} + \sqrt{\frac{2}{3}}\begin{array}{|c|c|c|}\hline & & \\ \hline & \downarrow & \\ \hline\end{array}$ and $\left|j = 3/2,\ m_j = -3/2\right\rangle = \begin{array}{|c|c|c|}\hline & & \\ \hline \downarrow & & \\ \hline\end{array}$.

In practice, when we measure the *total* angular momentum $|\mathbf{J}|^2$ of a one-electron atom, then each multiplet of states of given $l > 0$ yields *two* groups of states characterized by two values of j, namely $l - 1/2$ and $l + 1/2$. Unless spherical symmetry is broken, all states at given j and different m_j have the same energy. Even states of different j are degenerate as long relativistic effects are ignored, but no physical reason prevents them to acquire different energies. In Sect. 2.1.7 we discuss a weak relativistic interaction associating slightly different energies to different j, thus clarifying the need for the coupled basis. This interaction provides an explanation for the observed *two*-fold fine-split structure of the spectral lines (Fig. 2.3).

The angular-momentum composition rules discussed here are more general than the composition of the orbital and spin angular momenta of a single electron. These rules are purely algebraic: they apply equally well to any kind of angular momentum. We rely on this formalism for combining the angular momenta of many-electron atoms, especially in Sects. 2.2.4 and 2.2.8.3.

B.8.2 Coupled Magnetic Moments

The magnetic properties of a rotating charge are determined by angular-momentum properties, Eq. (2.56). In this Section we evaluate the matrix elements of the magnetic-moment operator on the spin-orbital uncoupled and coupled basis. These matrix elements are related *via* the unitary transformation (B.73). In the uncoupled basis $|m_l,\ m_s\rangle$, the μ_z operator is diagonal, with eigenvalues

$$\langle m_l,\ m_s|\mu_z|m_l,\ m_s\rangle = -\mu_{\mathrm{B}}\,\langle m_l,\ m_s|\frac{L_z + 2S_z}{\hbar}|m_l,\ m_s\rangle = -\mu_{\mathrm{B}}\,(m_l + 2m_s)\,. \tag{B.74}$$

The matrix elements of the μ_x and μ_y components too have explicit expressions, which can be obtained from the well-known (nondiagonal) matrix elements of $L_{x/y}$ and $S_{x/y}$ [17].

In principle one could obtain the matrix elements of $\boldsymbol{\mu}$ in the coupled basis $|j,\ m_j\rangle$ by using explicitly the Clebsch-Gordan transformation (B.73). However, a simpler and more instructive method yields these matrix elements within each subspace at fixed total angular momentum j. The key point is a symmetry argument: on average, *all vector quantities* characterizing a spherically symmetric object freely rotating in space are *proportional to its total angular momentum*. This means in particular that

$$\langle\boldsymbol{\mu}\rangle \propto \langle\mathbf{J}\rangle, \quad \langle\mathbf{L}\rangle \propto \langle\mathbf{J}\rangle, \quad \text{and} \quad \langle\mathbf{S}\rangle \propto \langle\mathbf{J}\rangle\,.$$

The matrix elements of **J** are well known. The only unknown quantities are the individual proportionality constants. To obtain the relevant ones, note that by definition

$$\begin{aligned}\langle j,\, m_j|\boldsymbol{\mu}|j,\, m_j\rangle &= -\frac{\mu_\mathrm{B}}{\hbar}\,\langle j,\, m_j|\mathbf{L}+2\mathbf{S}|j,\, m_j\rangle = -\frac{\mu_\mathrm{B}}{\hbar}\,\langle j,\, m_j|\mathbf{J}+\mathbf{S}|j,\, m_j\rangle \\ &= -(1+\gamma)\,\frac{\mu_\mathrm{B}}{\hbar}\,\langle j,\, m_j|\mathbf{J}|j,\, m_j\rangle = -g_j\,\frac{\mu_\mathrm{B}}{\hbar}\,\langle j,\, m_j|\mathbf{J}|j,\, m_j\rangle\,,\end{aligned} \tag{B.75}$$

where we introduce the ratio γ between $\langle j,\, m_j|\mathbf{S}|j,\, m_j\rangle$ and $\langle j,\, m_j|\mathbf{J}|j,\, m_j\rangle$. We determine γ by observing that the same ratio is involved when we take the scalar product with **J**:

$$\begin{aligned}\langle j,\, m_j|\mathbf{S}|j,\, m_j\rangle &= \gamma\,\langle j,\, m_j|\mathbf{J}|j,\, m_j\rangle\,, \\ \langle j,\, m_j|\mathbf{J}\cdot\mathbf{S}|j,\, m_j\rangle &= \gamma\,\langle j,\, m_j|\mathbf{J}\cdot\mathbf{J}|j,\, m_j\rangle\,.\end{aligned} \tag{B.76}$$

γ can be extracted from Eq. (B.76) by replacing the scalar product $\mathbf{J}\cdot\mathbf{S}$ with the expression

$$\mathbf{J}\cdot\mathbf{S} = \frac{|\mathbf{J}|^2+|\mathbf{S}|^2-|\mathbf{L}|^2}{2}\,, \tag{B.77}$$

obtained by squaring $(\mathbf{J}-\mathbf{S}) = \mathbf{L}$. [Note the similarity to Eq. (2.36).] We obtain the proportionality constant

$$\begin{aligned}\gamma &= \frac{\langle j,\, m_j|\mathbf{J}\cdot\mathbf{S}|j,\, m_j\rangle}{\langle j,\, m_j|\,|\mathbf{J}|^2\,|j,\, m_j\rangle} = \frac{\langle j,\, m_j|\,|\mathbf{J}|^2+|\mathbf{S}|^2-|\mathbf{L}|^2\,|j,\, m_j\rangle}{2\,j(j+1)\,\hbar^2} \\ &= \frac{j(j+1)+s(s+1)-l(l+1)}{2j(j+1)}\,.\end{aligned}$$

This result allows us to evaluate the proportionality constant between $\boldsymbol{\mu}/\mu_\mathrm{B}$ and $-\mathbf{J}/\hbar$ introduced in Eq. (B.75):

$$g_j = 1+\gamma = 1+\frac{j(j+1)+s(s+1)-l(l+1)}{2j(j+1)} \tag{B.78}$$

called *Landé g-factor*. g_j measures (in units of μ_B) the *effective* atomic magnetic moment resulting from the combined orbital and spin contributions, as seen within a given fixed-j multiplet. In the spectroscopy of atoms the states of such a j-multiplet split in a magnetic field as if they had a component of $\langle\boldsymbol{\mu}\rangle$

$$\langle j,\, m_j|\mu_z|j,\, m_j\rangle = -g_j\,\mu_\mathrm{B}\,m_j \tag{B.79}$$

in the $\hat{\mathbf{z}}$ direction of the field. Be warned that not all off-diagonal matrix elements of S_z (and thus of μ_z) vanish in the coupled basis $|j, m_j\rangle$. Note also that contrary to what one might expect for a combination of two magnetic moments with $g_l = 1$ and $g_s = 2$, the values of g_j are *not* restricted to the range $1 \le g_j \le 2$.

B.9 Perturbation Theory

Real QM system are occasionally similar to simple systems for which exact analytic solutions are available, like in the examples of Sect. B.7. This similarity is expressed by

$$H = H_0 + V\,, \tag{B.80}$$

where H_0 is the "simple-system" Hamiltonian, and the operator $V = H - H_0$ is the difference between the full Hamiltonian for the actual problem and the simple Hamiltonian. Consider now

$$H^{(\lambda)} = H_0 + \lambda\, V\,, \tag{B.81}$$

where $0 \le \lambda \le 1$ is a real parameter which tunes continuously the Hamiltonian system from the simple H_0 to the full H. If V is sufficiently "small" (meaning that H_0 is really similar to H), then the set of eigenenergies and eigenkets should evolve continuously as follows:

$$\begin{array}{rcl} \lambda & : & 0 \to 1 \\ H^{(\lambda)} & : & H_0 \to H \\ E_n^{(\lambda)} & : & E_n^{(0)} \to E_n \\ |n^{(\lambda)}\rangle & : & |n^{(0)}\rangle \to |n\rangle\,. \end{array} \tag{B.82}$$

It should be possible to Taylor-expand the functional dependence of the eigenenergy and eigenkets:

$$E_n^{(\lambda)} = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots \tag{B.83}$$

$$|n^{(\lambda)}\rangle = |n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \ldots \tag{B.84}$$

It is possible to prove that the first-order energy correction

$$E_n^{(1)} = \langle n^{(0)}|V|n^{(0)}\rangle\,. \tag{B.85}$$

The second-order energy correction is

$$E_n^{(2)} = \sum_{k \neq n} \frac{|\langle n^{(0)}|V|k^{(0)}\rangle|^2}{E_n^{(0)} - E_k^{(0)}} \,. \tag{B.86}$$

The first-order correction to the eigenkets is

$$|n^{(1)}\rangle = \sum_{k \neq n} |k^{(0)}\rangle \, \frac{\langle n^{(0)}|V|k^{(0)}\rangle}{E_n^{(0)} - E_k^{(0)}} \,. \tag{B.87}$$

The corrections of Eqs. (B.85)–(B.87) are used to generate approximate eigenvalues and eigenkets by substitution into Eqs. (B.83) and (B.84), taking $\lambda = 1$. For example, the first-order approximation to the eigenvalues of H is given by

$$E_n \simeq E_n^{(0)} + E_n^{(1)} = E_n^{(0)} + \langle n^{(0)}|V|n^{(0)}\rangle \,. \tag{B.88}$$

Note that:

- Expression (B.85)–(B.88) involve *unperturbed* energies and matrix elements of the perturbing Hamiltonian V over *unperturbed* states. All these quantities involving the simple system H_0 can either be evaluated as analytic expressions. or be computed by numerical integration.
- Equations (B.86) and (B.87) involve energy denominators $E_n^{(0)} - E_k^{(0)}$ which vanish for degenerate H_0-eigenstates. In such degenerate cases, Eqs. (B.86) and (B.87) do not hold as such. However, in a degenerate situation one is free to chose any linear combination of the unperturbed basis states within the degenerate space. One can then select suitable combinations such that all off-diagonal matrix elements $\langle n^{(0)}|V|k^{(0)}\rangle$ at the numerators above the vanishing denominators vanish as well. This means that one must first diagonalize the perturbation operator within all degenerate subspaces of the H_0 system. After this diagonalization, Eqs. (B.86) and (B.87) hold again, with the provision that the $0/0$ terms are set to 0.
- The kets in any truncated version of Eq. (B.84) are generally unnormalized. It is straightforward to normalize each approximate eigenket by dividing it by its norm.

An important consequence of Eq. (B.85) is that perturbing *operators with vanishing diagonal matrix elements induce null first-order corrections* to the eigenenergies. The lowest nonzero correction is then second order and involves the ratio of the square of off-diagonal matrix elements divided by the unperturbed energy difference, Eq. (B.86). This quadratic dependence often results in *very* small corrections. For example, if $E_n^{(0)} - E_k^{(0)} \simeq 1$ eV and $\langle n^{(0)}|V|k^{(0)}\rangle \simeq 1$ meV, then the lowest-order correction is $E_n^{(2)} \simeq (10^{-3}\text{ eV})^2/(1\text{ eV}) \simeq 1\ \mu\text{eV}$.

An apparent exception to this observation occurs in the degenerate case $E_n^{(0)} - E_k^{(0)} = 0$. As noted above, this singular condition must be treated by pre-diagonalizing the perturbation operator in the degenerate space. After this unitary rotation, the perturbation V becomes purely diagonal: Eq. (B.85) applies, and predicts first-order corrections, of the same order as the perturbation matrix elements.

B.10 Interaction of Charged Particles and Electromagnetic Radiation

A charged particle such an electron or a proton interacts with the electromagnetic fields. The form of this interaction is

$$H = \frac{1}{2m} (\mathbf{P} - q\mathbf{A})^2 + q\phi + V \,, \tag{B.89}$$

where and m, q are the particle's mass and charge, $\mathbf{A}$ and ϕ are the electromagnetic vector and scalar potentials and V is the potential energy describing non-electromagnetic forces acting on the particle. The potentials are functions of position (and often of time). In QM the particle position $\mathbf{R}$ is an operator. Therefore the electromagnetic potentials are operators, and precisely they are functions of the position operator $\mathbf{R}$. This implies that while $[A_u(\mathbf{R}), R_v] = 0$, in general $[A_u(\mathbf{R}), P_v] \neq 0$, so the potentials and the particle momentum are not compatible observables. The meaning of the square in Eq. (B.89) is then

$$(\mathbf{P} - q\mathbf{A})^2 = |\mathbf{P}|^2 - q(\mathbf{P} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{P}) + q^2 |\mathbf{A}|^2 \,. \tag{B.90}$$

Of these three terms, the first one generates the standard kinetic energy of the uncoupled particle, the others describe the particle-fields coupling. For static fields, one can represent the electric fields in terms of the ϕ potential, and the magnetic field in terms of $\mathbf{A}$. When the fields change in time, as in the presence of electromagnetic radiation, a suitable gauge choice must be adopted.

As long as the electromagnetic fields are not too intense, this interaction is usually addressed by *time-dependent perturbation theory*, a rather advanced topic whose mathematical subtleties far exceed the level of the present Appendix. To our purposes it suffices to retain that the electric-dipole emission rate Eq. (2.45) is obtained by a linear-response perturbative treatment of the term proportional to q in Eq. (B.90). The scattering of radiation, e.g. in the investigation of the structure of solids as in Sect. 5.1.3, in terms of Eq. (B.90) is described as a first-order contribution of the q^2 term plus the second-order effect of the q term.

Solutions

Problems of Chap. 2

2.1 5 components; 25.4 mm
2.2 $1s^2 2s^2 2p^5 3s\ {}^1P_1$, $1s^2 2s^2 2p^5 4s\ {}^1P_1$; both split into 3 components separated by 58 μeV
2.3 5145 eV; 2827 eV, 26835 eV, 31199 eV
2.4 $\gamma_{3s-1s} = 0$; $\gamma_{2p-1s} = 6.27 \times 10^8\ s^{-1}$
2.5 7.177×10^{-23} A m^2; 5 components
2.6 4 lines; 3 sub-lines for each line with $\Delta J = 0$, 6 sub-lines for the lines $\Delta J = 1$ and $\Delta J = -1$
2.7 ${}^3G_{3/4/5}$, ${}^3F_{2/3/4}$, ${}^3D_{1/2/3}$; ${}^3G_3$, ${}^3F_3$, ${}^3D_3$
2.8 1.606×10^{-23} A m^2; 1.606×10^{-23} A m^2; 5.487×10^{-23} A m^2
2.9 24350 eV (K); 3604 eV (L_I); 3330 eV (L_{II}); 3173 eV (L_{III}); the K shell; 42.30
2.10 21.25 %; 0.18°
2.11 L, 5400 eV; M, 1000 eV
2.12 $\frac{2}{5} E_{\mathrm{Ha}} (r_m/a_0)^2 = 3.15$ neV

Problems of Chap. 3

3.1 $l = 4$; 2581.3 cm^{-1} and 2735.9 cm^{-1}
3.2 4.086×10^{-10} N; 0.0183 J mol^{-1} K^{-1}
3.3 380 kg/s; 1.46×10^{-10} m; 280 K
3.4 14 levels; $Z = 3.885$; bound fraction = 97.25 %
3.5 2.297×10^{-9} N; 1.508 meV
3.6 62964 GHz; 50.6 meV (in the harmonic approximation)
3.7 21.459 J mol^{-1} K^{-1}
3.8 1.3575×10^{-18} J; 1.0134×10^{-10} m
3.9 9.42110^{13} Hz

N. Manini, *Introduction to the Physics of Matter*,
Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-14382-8

3.10 $l = 4$ [provided that one takes into account all the following effects: (i) the decrease of the zero-point energy, (ii) the increase of V_{ad}, and (iii) the decrease of the V_{centrif} term. $l = 8$ in case effects (i) and (ii) were ignored]
3.11 74.08 pm; 6.19 pm; 0.084
3.12 343.7 nm

Problems of Chap. 4

4.1 $16\pi^2\hbar\nu^3c^{-3}\rho_\nu^{-1} = 5.297 \times 10^{-21}$
4.2 $P = \hbar c\,(N/V)^{4/3}(3\pi^2)^{1/3}/4 = 5.712 \times 10^{24}$ Pa
4.3 1.401×10^{10} Pa; 1.633×10^{6} Pa
4.4 percentage ($J = 1/2$) = 30.81 %; 2.07×10^{-23} J/K; 3.963×10^{-26} J/K
4.5 1.70
4.6 0.42102 J mol^{-1} K^{-1}; 0.0327
4.7 1.372×10^{-6} J/(kg K)
4.8 15.625; 39.06
4.9 1.046×10^{6} m/s; 2.090×10^{9} Pa
4.10 2046 K; 14.372
4.11 8953 K
4.12 944 W; 302.1 K
4.13 4.96 GPa; $\Delta P = 1.76$ GPa
4.14 8.0 mW

Problems of Chap. 5

5.1 1.519×10^{6} m/s
5.2 $\alpha = 3$; $A = 1.831 \times 10^{-5}$ J mol^{-1} K^{-4}; 23.9 K
5.3 427.5 nm
5.4 gap $\simeq E = 6.8$ eV; insulator; transparent
5.5 A: 512.6 pm; B: 887.8 pm; C: 724.9 pm; A, with density 1479 kg m^{-3}
5.6 11110 m/s
5.7 $\omega(k) = 2\sqrt{C/M_{\mathrm{Fe}}}\,\sin(ka/2)$; 8.26 THz
5.8 20.7 μs
5.9 24.35°
5.10 $\Theta_D = 360$ K; $v_s = 3420$ m/s; approximately $6R \simeq 50$ J mol^{-1} K^{-1}

References

1. R. Eisberg, R. Resnick, *Quantum Physics* (Wiley, New York, 1974)
2. J.J. Brehm, W.J. Mullin, *Introduction to the Structure of Matter: A Course in Modern Physics* (Wiley, New York, 1989)
3. M. Alonso, E.J. Finn, *Fundamental University Physics: III. Quantum and Statistical Physics* (Addison-Wesley, Reading, 1968)
4. A.P. Arya, *Elementary Modern Physics* (Addison-Wesley, Reading, 1974)
5. E.U. Condon, H. Odabasi, *Atomic Structure* (Cambridge, New York, 1980)
6. B.H. Bransden, C.J. Joachain, *Physics of Atoms and Molecules* (Prentice Hall, Englewood Cliffs, 2003)
7. R.D. Cowan, *The Theory of Atomic Structure and Spectra* (University of California Press, Berkeley, 1981)
8. E.B. Wilson, J.C. Decius, P.C. Cross, *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra* (McGraw-Hill, New York, 1955)
9. G. Herzberg, *Molecular Spectra and Molecular Structure*, vol. III (Van Nostrand Reinhold, New York, 1966)
10. N.W. Ashcroft, M.D. Mermin, *Solid State Physics* (Holt-Saunders, Philadelphia, 1976)
11. C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1996)
12. G. Grosso, G.P. Parravicini, *Solid State Physics* (Academic, San Diego, 2000)
13. J.C. Slater, *Solid-State and Molecular Theory: A Scientific Biography* (Wiley, New York, 1975)
14. The web site http://physics.nist.gov/cuu/Constants/ provides a table of physical constants
15. L.I. Schiff, *Quantum Mechanics* (McGraw-Hill, Singapore, 1955)
16. P. Caldirola, R. Cirelli, G.M. Prosperi, *Introduzione alla Fisica Teorica* (UTET, Torino, 1982)
17. J.J. Sakurai, *Modern Quantum Mechanics* (Benjamin, Menlo Park, 1985)
18. T.W. Hänsch, I.S. Shahin, E.L. Schawlow, Nature Phys. Sci. **235**, 63 (1972)
19. B. Cordero, V. Gómez, A.E. Platero-Prats, M. Revés, J. Echeverría, E. Cremades, F. Barragán, S. Alvarez, Dalton Trans. **2008**, 2832 (2008)
20. The web site http://www.nist.gov/pml/data/asd.cfm provides a set of tables of atomic spectroscopy levels and transitions
21. J.C. Slater, *Quantum Theory of Matter* (McGraw-Hill, New York, 1951)
22. Several international facilities currently run sources of X-ray photons for research and applications. We list a few examples of web sites http://www.esrf.eu/, http://www.elettra.trieste.it/, http://www.spring8.or.jp/, http://www.diamond.ac.uk/, http://www-als.lbl.gov/, and http://www.bnl.gov/ps/
23. R.M. Dreizler, E.K.U. Gross, *Density Functional Theory: An Approach to the Quantum Many-Body Problem* (Springer, Berlin, 1990)
24. J. Kielkopf, K. Myneni, F. Tomkins, J. Phys. B: At. Mol. Opt. Phys. **23**, 251 (1990)
25. The web site http://materia.fisica.unimi.it/manini/dida/structures.html provides the 3D structures used in several illustrations of the present textbook

N. Manini, *Introduction to the Physics of Matter*,
Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-14382-8

26. Z.C. Yan, J.F. Babb, A. Dalgarno, G.W.F. Drake, Phys. Rev. A **54**, 2824 (1996)
27. J.L. Gardner, J.A.R. Samson, J. Chem. Phys. **62**, 1447 (1975)
28. R.E. Grisenti, W. Schöllkopf, J.P. Toennies, G.C. Hegerfeldt, T. Köhler, M. Stoll, Phys. Rev. Lett. **85**, 2284 (2000)
29. G.H. Wannier, *Statistical Physics* (Wiley, New York, 1966)
30. R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics* (Wiley, New York, 1975)
31. G. Turrell, *Mathematics for Chemistry and Physics* (Academic, London, 2002)
32. W.H. Lien, N.E. Phillips, Phys. Rev. **133**, A1370 (1964)
33. H. Kleinke, Eur. J. Inorg. Chem. **1998**, 1369 (1998)
34. L.P. Pitaevskii, S. Stringari, *Bose-Einstein Condensation* (Oxford University Press, Oxford, 2003)
35. R. Bini, L. Ulivi, H.J. Jodl, P.R. Salvi, J. Chem. Phys. **103**, 1353 (1995)
36. The web site http://www.lks.physik.uni-erlangen.de/diffraction/ provides a practical interactive tool for diffraction simulation
37. T. Proffen, R.B. Neder, S.J.L. Billinge, J. Appl. Cryst. **34**, 767 (2001)
38. K. Koyama, T. Kajitani, Y. Morii, H. Fujii, M. Akayama, Phys. Rev. B **55**, 11414 (1997)
39. E.O. Wollan, C.G. Shull, Phys. Rev. **73**, 830 (1948)
40. H.J. Fritzsche, J. Phys. Chem. Solids **6**, 69 (1958)
41. The web site http://britneyspears.ac/lasers.htm maintains a detailed and entertaining “hands on” approach to semiconductors physics and technology (beware of a few minor mistakes!)
42. F.M.F. de Groot, M.H. Krisch, J. Vogel, Phys. Rev. B **66**, 195112 (2002)
43. E.D. Palik, *Handbook of Optical Constants of Solids* (Academic, San Diego, 1998)
44. M.A. Green, Sol. Energ. Mat. Sol. Cells **92**, 1305 (2008)
45. G. Onida, L. Reining, A. Rubio, Rev. Mod. Phys. **74**, 601 (2002)
46. K.K. Rao, T.J. Moravec, J.C. Rife, R.N. Dexter, Phys. Rev. B **12**, 5937 (1975)
47. D.W. Berreman, Phys. Rev. **130**, 2193 (1963)
48. J.P. Elliott, P.G. Dawber, *Symmetry in Physics* (McMillan, London, 1979)
49. P.H. Butler, *Point Group Symmetry Applications* (Plenum, New York, 1981)
50. A.R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, 1974)